CODEN: 4966

THE JOURNAL of the Acoustical Society of America

Vol. 102, No. 2, Pt. 1 August 1997

(Continued)

CONTENTS—*Continued from preceding page*

(Continued)

CONTENTS—*Continued from preceding page*

(Continued)

CONTENTS—*Continued from preceding page*

NOTES CONCERNING ARTICLE ABSTRACTS

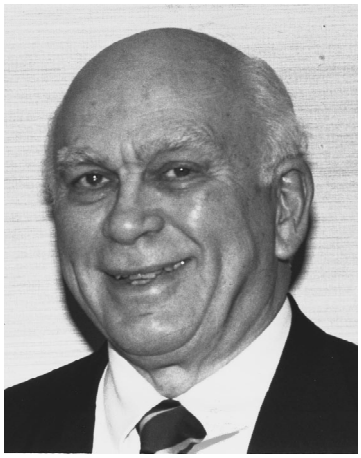
1. The number following the abstract copyright notice is a Publisher Item Identifier (PII) code that provides a unique and concise identification of each individual published document. This PII number should be included in all document delivery requests for copies of the article.
2. PACS numbers are 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.

ACOUSTICAL NEWS—USA

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

Editor's Note: *Deadline dates for news items and notices are 2 months prior to publication.*

New Fellows of the Acoustical Society of America



R. Craig Olson—For contributions to underwater acoustics.

Teruji Yamamoto—For contributions to the acoustics of performance spaces.

Appreciation to the 1996 reviewers of manuscripts submitted to the *Journal*

The quality and dependability of the information contained in the articles and letters published in the *Journal of the Acoustical Society of America* are assured through the competent peer reviewing contributed by the reviewers that our Associate Editors enlist each year. The reviews show evidence of the reviewers' diligence and dedication to the Society and to their profession. In appreciation to the reviewers for these anonymous services to the *Journal* and its authors during 1996, the Editor-in-Chief gratefully publishes their names in alphabetical order without identification of their review specialities or of the articles that they reviewed. In such a long list (1479) some errors and omissions are certain to occur. If you reviewed any *Journal* article or letter in 1996 but your name has been omitted (or misspelled), please accept our apology and send a correction promptly to the Editor-in-Chief. Thank you.

DANIEL W. MARTIN
Editor-in-Chief

1996 JASA Peer Reviewers

Abawi, A.
Abbas, P. J.
Abdala, C.
Abel, S.
Accorsi, M. L.
Ackerman, H.
Adams, S. G.
Adler, L.
Agullo, J.
Ainslie, M.
Akahane-Yamada, R.
Akay, A.
Albert, D. G.
Alcantara, J. I.
Alipour, F.
Alku, P.
Allard, J.-F.
Allen, J. B.
Allen, P.
Anderson, A.
Anderson, M. J.
Andrew, R.
Andrews, M.
Apfel, R. E.
Arehart, K.
Arnott, W. P.
Ashmore, J. F.
Askenfelt, A. G.
Aslin, R. N.
Assman, P.
Astor, M.
Atchley, A. A.
Attala, N.
Attenborough, K.
Au, W. W. L.
Augspurger, G. L.
Avan, P.
Ayers, G.
Ayers, R. D.
Ayers, T. J.
Bachman, R.
Bacon, S.
Badin, P.
Baer, T.
Bailey, P. J.
Baken, R. J.
Baker, A. C.
Baker, C. R.
Ballandras, S.
Ballato, A.
Barber, B. P.
Barbone, P. E.
Barger, J. E.
Bargones, J. Y.
Barin, K.
Barron, M.
Barry, W. J.
Barthé, P. G.
Bartram, J. F.
Bass, H. E.
Baumann, W. T.
Bautista, E.
Beckman, M.
Beitel, R. E.
Bell, A. E.
Bell-Berti, F.
Benkeser, P.
Bennett, S.
Beran, M. J.
Beranek, L. L.
Berg, B.
Berg, K. M.
Berge, T.
Berglund, B.
Berke, G. S.
Berman, D. H.
Bernard, M.
Bernhard, R. J.
Bernstein, J. J.
Bernstein, L.
Berry, A. C.
Berry, D. A.
Bert, C. W.
Berthelot, Y. H.
Best, C.
Beyer, R. T.
Bhat, R. B.
Bianco, B.
Bickley, C.
Bielak, J.
Bilger, R. C.
Bilsen, F.
Bird, J. F.
Birdsall, T. G.
Bishop, G. C.
Bissinger, G. A.
Blake, W. K.
Blamey, P.
Blauert, I. J.
Blauert, J. P.
Bloothoft, G.
Boë, L-J.
Boettcher, F. A.
Bohn, O-S.

 0001-4966/97/102(2)/663/10/$10.00

Bond, L. J.
Bonfils, P.
Bongiovanni, K.
Boone, M. M.
Booth, N.
Boothroyd, A.
Borgiotti, G. V.
Bork, I.
Boström, A.
Bottega, W. J.
Botteldooren, D. B.
Bouchard, M.
Boucher, C. C.
Boucher, D.
Boutillon, X.
Bowen, D. L.
Boyce, S. E.
Boyle, F. A.
Boyles, C. A.
Bradley, C. E.
Bradley, J. S.
Bradlow, A.
Braida, L.
Brass, D.
Breazeale, M. A.
Bregman, A.
Brennan, M. J.
Brewster, J.
Briers, R.
Briggs, G. A. D.
Brill, D. W.
Brooke, G. H.
Broschat, S. L.
Brown, A. M.
Brown, C. J.
Brown, D. A.
Brown, J.
Brown, J. C.
Brown, M. G.
Brown, W. S., Jr.
Brownell, W. E.
Bruce, G.
Brudner, T. J.
Bryne, C. L.
Bucaro, J. A.
Bucker, H.
Buckingham, M. J.
Buckley, K. M.
Buell, T. N.
Burdisso, R. A.
Burkard, R. F.
Burke, J. E.
Burnett, D. S.
Burns, E.
Burroughs, C. B.
Busby, P. A.
Busch-Vishniac, I. J.
Butler, J. L.
Butler, R. A.
Buus, S.
Byrne, D.
Cacace, A. T.
Cahn, J.
Caiti, A.
Calabrese, A.
Campanella, A.
Candy, J. V.
Canevet, G. L.
Cantrell, J. H.
Carbo, R.
Carcione, J. M.
Carey, W. M.
Carlyon, R. P.
Carneal, J. P.
Carney, A.
Carney, L. H.
Carr, C. E.
Carre, R.
Carrell, T.
Carson, P. L.
Carstens, B.
Carstensen, E. L.
Caruso, A. J.
Caspary, D. M.
Cassereau, D.
Castagnede, B.
Castelli, E.
Cato, D. H.
Cawley, P.
Ceperley, P. H.
Chai, J-F.
Chaigne, A. J.
Chakraborty, B.
Chambers, J.
Chamlin, C.
Champoux, Y.
Chamuel, J. R.
Chan, M. K.
Chandiramani, K. L.
Chapman, C. H.
Chapman, D. M. F.
Chapman, R.
Charette, F.
Chase, D. M.
Chattopadhyay, A.
Cheeke, J. D. N.
Cheenne, D.
Chen, E.
Chen, L.
Chen, P-T.
Chen, X.
Cheng, C. H.
Cheng, L.
Chiang, E.
Childers, D. G.
Chimenti, D. E.
Chin-Bing, S. A.
Chisolm, T.
Chiu, C-S.
Cho, Z. H.
Cholewiak, R. W.
Chotiros, N. P.
Chou, W.
Chu, D.
Chu, W. T.
Chu, Y. C.
Church, C. C.
Ciskoski, R. D.
Clark, M., Jr.
Clark, W. W.
Clay, C. S.
Cleveland, R.
Clopton, B. M.
Cochrane, N.
Cody, A.
Cohen, A.
Cohen, Y. E.
Coleman, A. J.
Collins, J. A.
Collins, L.
Collins, M. D.
Collins, M. J.
Colosi, J.
Coltman, J. W.
Colton, D.
Commander, K. W.
Comparin, R. J.
Comte-Bellot, G.
Conklin, H. A.
Conoir, J. M.
Cook, P. R.
Cooke, M.
Cops, A.
Cornuelle, B.
Cotaras, F. D.
Couchman, L. S.
Cowan, R. S. C.
Cox, D.
Craig, J. C.
Cranen, B.
Craster, R. V.
Crawford, A. M.
Cray, B. A.
Crouther, C.
Crum, L. A.
Cruttenden, A.
Cullen, P. J.
Culling, J. F.
Cummings, A.
Cummings, W. C.
Cunefare, K. A.
Cuomo, F. W.
Cuschieri, J. M.
Cywiak, M.
Dabirikhah, H.
Dacol, D.
Dahl, P. H.
Dai, H.
Daigle, G. A.
Dalebout, S.
Dalecki, D.
Dalenback, B-I.
d'Alessandro, C. R.
Dancer, A. L.
Dang, J.
Dannenberg, R. B.
Darling, T. W.
Darlington, P.
Daroux, P. A.
Darvennes, C.
Darwin, C. J.
Dassios, G.
Dau, T.
David, P. M.
Davies, H.
Davies, P.
Davis, L.
Dawson, T. W.
Dawson, T. W.
de Billy, M.
de Boer, B.
de Boer, E.
de Cheveigne, A.
de Jong, K.
de Krom, G.
de Moustier, C. P.
Deane, G. B.
DeFilippo, C.
Dehandschutter, W.
DeJong, K.
DeJong, R. G.
Delcomyn, F.
Delgutte, B.
Demany, L. J. D.
Demer, D. A.
de Pijper, J. R.
Demorest, M. E.
Deng, L.
Deutsch, D.
Devaney, A. J.
Dey, S.
Di Iorio, D.
Di, X.
Dickey, J. W.
Diebold, G. J.
Diehl, R.
Dimaragonas, A. D.
Ding, L.
Dirks, D.
Dixon, T. L.
Docherty, G.
Doherty, K.
Doll, T. J.
Don, C.
Don, M.
Dooling, R. J.
Dorman, M.
Dosso, S. E.
Doucet, J. R.
Douglas, B. L.
Dowell, E. H.
Dowling, J. P.
Dozier, L.
Dragonette, L.
Drullman, R.
Dubbleday, P. S.
Dubno, J.
Dubno, J. R.
Duda, T. F.
Duifhuis, H.
Dunn, F.
Dushaw, B. D.
Dutoit, T.
Dye, R. H., Jr.
Dywer, R. F.
Ebbini, E.
Ebenezer, D. D.
Eddins, A.
Eddins, D.
Edmonds, P. D.
Edwards, B.
Edwards, J.
Eggermont, J. J.
Eisenberg, L.
Eisler, T.
El-Raheb, M.
Elko, G. W.
Eller, A. I.
Ellis, D.
Embleton, T. F. W.
Eriksson, L. J.
Espy-Wilson, C. Y.
Estill, J.
Evans, R. B.
Evans, W. E.
Everbach, E. C.
Eversman, W.
Everstine, G. C.
Every, A. G.
Ewart, T. E.
Faber A.
Fabry, D. A.
Fahey, P. F.
Fahnline, J. B.

Fahy, F. J.
Fant, G.
Fantini, D.
Farag, H. H.
Fastl, H.
Faulkner, A.
Fawcett, J. A.
Fay, R. R.
Feit, D.
Feng, L.
Ferguson, B. G.
Festen, J.
Feuillade, C.
Ffowcs Williams, J. E.
Fidell, S.
Filipczynski, L.
Filippi, P.
Finch, R. D.
Finette, S.
Finley, C.
Fishman, L.
Fitzgibbons, P.
Flatté, S. M.
Flege, J. E.
Fletcher, J. M.
Fletcher, N. H.
Florentine, M.
Foote, K. G.
Ford, R.
Formby, C.
Forrest, T. G.
Fourakis, M.
Fowler, C. A.
Fowlkes, J. B.
Fox, R.
Frank, T. A.
Franke, S. J.
Frankenthal, S.
Franks, J. R.
Franzoni, L. P.
Freedman, A.
Freyman, R. L.
Frisina, D. R., Sr.
Frizzell, L. A.
Fu, S.-S.
Fuller, C. R.
Funnell, W. R. J.
Fyfe, K.
Gadd, C.
Gade, A. C. S.
Gaitan, D. F.
Galindo, M.
Gallego-Juarez, J. A.
Galloway, W. J.
Garcia-Bonito, J.
Garrelick, J. M.
Garrett, S.
Gauffin, J.
Gaumond, C. F.
Gaumond, R. P.
Gaunaurd, G. C.
Gauss, R.
Ge, H.
Ge, L-F.
Geddes, E. R.
Geisler, C. D.
Genuit, K.
Georgiadis, H. G.
Gerken, L-A.
Gerratt, B. R.
Gerretson, E.
Gerstoft, P.
Gescheider, G. A.
Gibson, R. L.
Gilbert, K. E.
Gilkey, R. H.
Gillis, K. A.
Gimenez, G.
Gingold, H.
Ginsberg, J. H.
Gitza, O.
Glasberg, B.
Glattke, T. J.
Glegg, S.
Gleich, B.
Godin, O. A.
Goldinger, S. D.
Goldstein, J. L.
Goode, R. L.
Goodman, R.
Gordon, P.
Gordon-Salant, S.
Gorga, M. P.
Goss, S. A.
Gottfried, T.
Gough, G. O.
Goumilevski, A. G.
Grabb, M.
Graham, W. H.
Graham, W. R.
Grantham, W.
Grattan, K. T. V.
Gray, L.
Greaves, R. J.
Green, D. M.
Green, K.
Greenberg, J. E.
Greenberg, S.
Greenlaw, C. F.
Griffin, D. R.
Grose, J.
Grosh, K.
Güdeson, A.
Guigou, C.
Guo, Y.
Guo, Y. P.
Gustafson, G.
Guy, R. W.
Guyader, J.-L.
Guzman, S. J.
Haddadin, O. S.
Hafter, E.
Hall, D. E.
Hall, J. L.
Hall, J. W.
Hall, M. V.
Hall, T.
Hallé, P. A.
Hallworth, R.
Hambric, S. A.
Hamernik, R. P.
Hamilton, M. F.
Hammarberg, B.
Handel, S.
Hanna, T. E.
Hansen, C.
Hansen, T. B.
Hansen, U. J.
Hanson, H. M.
Hardcastle, P.
Hardin, J. C.
Harding, A. J.
Hari, R.
Harley, T.
Harrington, J.
Harris, G. R.
Harrison, C.
Hartmann, B.
Hartmann, W. M.
Hartung, K.
Havelock, D. I.
Hay, A. E.
Hayes, M.
Hayward, G.
He, N-J.
Heaney, K.
Heard, G. J.
Heatwole, C.
Hedberg, C. M.
Heffner, H. E.
Helfer, K. S.
Hellman, R.
Hellman, W.
Henderson, D.
Henton, C.
Herbert, R. K.
Hertegård, S.
Herzel, H.
Hewitt, M. J.
Heyman, E.
Hickling, R.
Hillenbrand, J.
Hillenbrand, J. M.
Hines, P. C.
Hirschberg, A.
Hixson, E. L.
Hodgson, M. R.
Hoffman, M. w.
Holford, R.
Holliday, D. V.
Hollins, M.
Holmberg, E.
Holt, R. G.
Holte, L. A.
Hood, L. J.
Hoole, P.
Horner, D.
Horst, J. W.
Hossack, J. A.
Hosten, B.
Hothersall, D. C.
House, A.
House, D.
Houtgast, T.
Houtsma, A. J.
Howard, D.
Howarth, T. R.
Howe, B. M.
Howe, M. S.
Howell, P.
Hua, Y.
Huang, D.
Huang, H.
Huang, W.
Hubbard, A. E.
Hudspeth, K. H.
Hughes, D. H.
Humes, L. E.
Hunt, H.
Hunten, D.
Huntley-Bahr, R.
Hutchins, D. A.
Hynynen, K.
Ianniello, J. P.
Idogawa, T.
Igusa, T.
Ih, J.-G.
Imaizumi, S.
Imhof, M. G.
Insana, M. F.
Isabelle, S. K.
Isakson, S.
Iverson, P.
Iwasa, K. H.
Izbicki, J-L.
Jackson, D. R.
Jackson, J. L.
Jacobsen, F.
Jacobson, M.
Jaffe, J. S.
Jaffee, J. C.
Jamieson, D. G.
Janata, P.
Jansson, E. V.
Jech, J. M.
Jenkins, J.
Jenkins, W.
Jensen, J. A.
Jerger, J. F.
Jesteadt, W.
Joharapurkar, D.
Johnson, D. H.
Johnson, D. L.
Johnson, K. A.
Johnson, K. O.
Johnson, L.
Johnson, P.
Johnson, R. A.
Johnson, S. A.
Jones, D. E.
Jones, D. F.
Jones, T. S.
Joyce, R. F.
Junger, M. C.
Kabal, P.
Kadambe, S.
Kaduchak, G.
Kalashnikov, A. N.
Kamakura, T.
Kampanis, N. A.
Kang, Y. J.
Karasalo, I.
Kargl, S. G.
Karlsson, I.
Karpur, P.
Kasuya, H.
Kay, K.
Keane, A. J.
Keating, P. A.
Keefe, D. H.
Keller, J.
Keltie, R. F.
Kerwin, E. M., Jr.
Ketten, D. R.
Kewley-Port, D.
Khdeir, A. A.
Kibblewhite, A. C.
Kidd, G. D.
Kidd, G. R.
Kim, D. O.
Kim, K. S.
Kim, M.
Kim, Y-H.
Kimball, C. V.

Kimberley, B. P.
Kindel, J.
King, M. E.
Kinra, V. K.
Kistler, D.
Klauson, A.
Kleiner, M.
Kluender, K.
Knobles, D.
Ko, S-H.
Koch, R. A.
Koehnke, J.
Kohlrausch, A.
Kolaini, A. R.
Kollmeier, B.
Koopmann, G. H.
Kossl, M.
Krakow, R. A.
Kraman, S. S.
Krane, M. H.
Kreiman, J.
Kriegsmann, G. A.
Kringlebotn, M.
Krishnamurthy, A. K.
Krothapalli, A.
Krull, D.
Kubota, H.
Kuc, R.
Kuhl, P. K.
Kumar, A.
Kundu, T.
Kuperman, W. A.
Kurze, U.
Kuttruff, H.
Kuwada, S.
L'Esperance, A.
Ladd, D. R.
Lafleur, L. D.
Lafreniere, D.
Laible, H.
Lakhtakia, A.
Lalande, F.
Lam, C. F.
LaMancusa, J. S.
Langley, R. S.
Larraza, A.
Lataitis, R. J.
Lauchle, G. C.
Laugesen, S.
Laukkanen, A-M.
Laulagnet, B.
Laura, P. A. A.
Law, W.
Learned, R. E.
Lee, C. P.
Lee, H.
Lee, J.
Lee, M.-R.
Lee, Y. P.
Leek, M.
Leighton, T. G.
Lemond, J.
LePage, K. D.
Letcher, S. V.
Levitt, H.
Levy, M.
Lewis, E. R.
Lewis, K. D.
Lewis, N. D.
Lewis, T. J.
Li, S.
Li, Y. L.
Liang, Y-C.
Liberman, A. M.
Lieberman, P.
Liew, K. M.
Light, G. M.
Liljencrants, J. C. W. A.
Lim, R.
Lin, W.
Lisker, L.
Litovsky, R. Y.
Liu, Q-H.
Lizzi, F.
Ljolje, A.
Llemery, A.
Lo Vetri, J.
Long, G.
Longuet-Higgins, M.
Lonsbury-Martin, B. L.
Loughlin, P. J.
Love, R. H.
Lu, Y.
Lucas, R. J.
Lucero, J. C.
Ludwig, R.
Lupien, V.
Luschei, E. S.
Lutfi, R.
Lutgen, S. J.
Lutman, M.
Ly, L.
Lynch, J. F.
Lyon, R. H.
Lyons, A. P.
Ma, X.
Macaulay, M. C.
Mace, B. R.
MacLennan, D. N.
MacMillan, N.
Madanshetty, S.
Madden, J. P.
Madden, T.
Madigosky, W. M.
Madsen, E. L.
Maestrello, L.
Maidanik, G.
Maillard, J.
Makris, N. C.
Mal, A. K.
Mammano, F.
Maniatty, A.
Manis, P. B.
Manley, G.
Mann, A.
Mann, J. A., III
Manuel, S. Y.
Mao, Y.
Maranda, B.
Margolis, R. H.
Margulies, T. S.
Marin, C.
Maris, H.
Marler, P. R.
Marshall, L.
Marston, P.
Martin, D. W.
Martin, G. K.
Martin, S. A.
Martinez, R.
Mason, C.
Massaro, D.
Mast, T. D.
Masters, W.
Mattei, C.
Matula, T.
Maynard, J. D., Jr.
Mazur, M. A.
McAdams, S.
McCammon, D. F.
McCarthy, E.
McClements, D. J.
McCoy, J. J.
McDaniel, J. G.
McDevitt, T. E.
McDonald, B. E.
McEachern, J. F.
McFadden, D.
McFarland, D.
Meddis, R.
Medwin, H.
Meegan, G. D.
Mellert, V.
Meredith, R.
Mermelstein, P.
Meyer, J.
Michalopoulou, E. Z. H.
Middlebrooks, J. C.
Migliori, A.
Mignerey, P. C.
Mignogna, R. B.
Miki, N.
Miksis, M.
Milder, D. M.
Milios, E. E.
Miller, D. L.
Miller, E. L.
Miller, J. G.
Miller, J. H.
Miller, M. I.
Miller R. D.
Mills, D. M.
Mire, C. T.
Miskiewicz, A.
Moe, J. E.
Moffett, M. B.
Montgomery, A. A.
Moody, D. B.
Moon, S-J.
Moore, B. C. J.
Moore, C. A.
Moore, J. A.
Mooshammer, C.
Moss, C.
Mountain, D. C.
Moura, J. M. F.
Mourad, A.
Muir, T.
Mullennix, J.
Munhall, K.
Munjal, M. L.
Murakami, H.
Musicant, A. D.
Myer, L.
Myers, M. K.
Myrberg, A. A.
Nabelek, A.
Nachtigall, P. E.
Nagaya, K.
Naghshineh, K.
Nagy, P. B.
Nakayama, K.
Narendran, N.
Nashif, A. D.
Naugolnykh, K. A.
Nava, R.
Nayfeh, A. H.
Naz, P.
Nearey, T.
Nederveen, C. J.
Neely, S. T.
Neff, D.
Nelisse, H.
Nelson, D.
Nelson, P. A.
Newhouse, V. L.
Newman, J.
Nicolas, J.
Nightingale, T. R. T.
Nihei, K.
Nittrouer, S. N.
Noble, J. M.
Nolan, F.
Nolle, A. W.
Nord, L.
Nordebo, S.
Norris, A. N.
Norton, G. V.
Norton, S. J.
Norum, T.
Nossek, J. A.
Novarini, J. C.
Novotny, T.
Nowotny, H.
Nuttall, A. L.
Nyborg, W. L.
Nye, P. W.
Nystuen, J.
O'Brien, W. D., Jr.
O'Brien, W. O.
O'Donnell, M.
O'Keefe, J. P. M.
Ocheltree, K.
Ochmann, M.
Ogden, R. W.
Ogi, H.
Ohala, J. J.
Ohde, R.
Ohgushi, K.
Olson, J.
Onoe, M.
Ophir, J.
Orduna-Bustamante, F.
Orofino, D. P., II
Orris, G.
Ostrovskii, L.
Ostry, D. J.
Owsley, N.
Oxenham, A.
Ozisik, M. N.
Pace, N. G.
Palmer, A. R.
Palmer, D. R.
Pan, G. J.
Pan, J.
Papadakis, E.
Papazoglou, M.
Parastates, E.
Parham, K.
Park, Y. C.
Parker, K. J.
Parncutt, R.
Parra, J. O.
Parvulescu, A.

Pasterkamp, H.
Pastore, R.
Patterson, J. H., Jr.
Patterson, R. D.
Patton, T. C.
Patuzzi, R.
Paul, H. S.
Pavic, G.
Pawlowicz, R.
Payton, K. L.
Peake, W. T.
Pearson, K.
Pederson, P. C.
Pelissier, M-C.
Pelorson, X.
Peng, H.
Penner, M.
Penrose, J.
Perkell, J. S.
Perrott, D. R.
Peters, R. W.
Peterson, M.
Pfingst, B. E.
Pflug, L.
Phelps, A.
Photiadis, D. M.
Piacsek, A.
Pierce, A. D.
Pierce, J. R.
Piercy, J. E.
Pierucci, M.
Pind, J.
Pinkel, R.
Pinnington, R. J.
Piquette, J.
Pisoni, D.
Pitt, M.
Plack, C.
Plitnik, G. R.
Plomp, R.
Plotkin, E. I.
Polka, L.
Pollack, I.
Pollard, H. F.
Pols, L. C. W.
Ponton, C. W.
Popelka, G. R.
Popper, A. N.
Port, R. F.
Porter, R. J., Jr.
Potter, J. R.
Potty, G.
Povey, M.
Powell, A.
Powell, D. J.
Powell, R. E.
Powers, J. M.
Pressnitzer, D.
Preves, D.
Price, G. R.
Price, P. J.
Prieve, B. A.
Prijs, V. F.
Probst, R.
Prosen, C. A.
Prosperetti, A.
Prosser, W. H.
Pruitt, J.
Purcell, A.
Putterman, S.
Pyle, R. W., Jr.
Pyne, J. M.
Qatu, M. S.
Qi, Y.
Quinlan, D. A.
Quirt, J. D.
Raatgever, J.
Rabbitt, R. D.
Rabinowitz, W. M.
Radlinski, R. P.
Rahim, M.
Rajan, S. D.
Raju, P. K.
Rakerd, B. S.
Ramsay, G. J. T.
Randall, R.
Rankovic, C. M.
Rao, M. D.
Rapids, B.
Raspet, R.
Ravichandran, G.
Rawlins, A. D.
Razavy, M.
Readhead, M. L.
Rebinsky, D. A.
Recasens, D.
Reed, C.
Reid, J. M.
Relkin, E. M.
Remez, R.
Repp, B.
Revoile, S.
Reynolds, S.
Rhode, W. S.
Ribner, H. S.
Rice, D.
Richards, V. M.
Richardson, B. E.
Richardson, W. J.
Ricker, D. W.
Ricks, D. C.
Ridgway, S.
Rink, J.
Robertson, D.
Robertson, J. S.
Robertsson, J. O.
Robins, A. J.
Robinson, D. E.
Robinson, H. C.
Rochet, B.
Rodet, X.
Rodgers, O. E.
Rogers, J. C.
Rogers, P. H.
Rokhlin, S. I.
Rolt, K. D.
Rose, J. H.
Rose, J. L.
Rosen, S.
Rosowski, H.
Rosowski, J. J.
Rossetti, D. J.
Rossing, T. D.
Rossmanith, H. P.
Rouseff, D.
Roy, R.
Royster, J. D.
Royster, L. H.
Rubenstein, D.
Ruckman, C. E.
Ruggero, M. A.
Rumerman, M. L.
Russell, D. A.
Russell, G. A.
Russell, I. J.
Ryan, J. G.
Saberi, K.
Sachs, M. B.
Sachs, R. M.
Safaeinili, A.
Salama, K.
Saldana, H. M.
Salvi, R. J.
Samar, V. J.
Sammelmann, G. S.
Samuel, A.
Sanderson, M. A.
Sanes, D. H.
Sanghvi, N. T.
Santo-Sacchi, J. R.
Santossa, F.
Sarkar, K.
Sarkissian, A.
Sas, P.
Saunders, J. C.
Sawausch, J.
Scaife, R. P.
Scandrett, C.
Schaafsma, A. S.
Scharf, B.
Schechter, R. S.
Scheinbeim, J.
Schenck, H.
Scherer, K. R.
Scherer, R. C.
Schindall, J.
Schindel, D. W.
Schlauch, R.
Schmerr, L. W., Jr.
Schmidt, A.
Schmidt, H.
Schmidt-Nielson, A.
Schmiedt, R. A.
Schmitt, D. P.
Schneider, B.
Schoenberg, M.
Schoentgen, J.
Schomer, P.
Schouten, M. E. H.
Schroeder, M.
Schroeter, J.
Schuck, P. L.
Schulz, T.
Schum, D. J.
Schumacher, R. T.
Seiner, J. M.
Selamet, A.
Sessarego, J.-P.
Sethares, W. A.
Seybert, A. F.
Sgard, F.
Shackleton, T. M.
Shadle, C. H.
Shaiman, S.
Shamma, S. A.
Shang, E.-C.
Shankar, M. P.
Shannon, R. V.
Shapiro, S. A.
Sharma, A.
Sharp, B.
Shattuck-Hufnagel, S.
Shaw, E.
Sheft, S.
Shenderov, E. L.
Sheng, J.
Shera, C.
Sherman, C. H.
Shields, F. D.
Shigeno, S.
Shin, D. C.
Shinn-Cunningham, B. G.
Shofner, W. P.
Shooter, J. A.
Shung, K. K.
Siderius, M.
Siegel, J. H.
Siegmann, W. L.
Silva, A.
Simmons, J. A.
Sinex, D. G.
Singh, R.
Sinha, B. K.
Sininger, Y. S.
Sinnott, J.
Slaney, M.
Smith, C. C.
Smith, D. E.
Smith, E.
Smith, G.
Smith, J.
Smith, J. O., III
Smith, K. B.
Smith, L. K.
Smith, P. F.
Smith, R. L.
Smoorenburg, G. F.
Smurzynski, J.
Snell, K. B.
Snowdon, C.
Snyder, S. D.
Sock, R.
Soderman, P. T.
Soize, C.
Soli, S.
Sommerfeldt, S.
Sommers, M.
Sondhi, M. M.
Sorkin, R.
Sornette, D.
Sotirin, B.
Sparrow, V. W.
Spicer, J. B.
Spies, M.
Spiesberger, J. L.
Stack, J.
Stanic, S.
Stanton, T. K.
Stanzial, D.
Stebbins, W. C.
Steinberg, B. D.
Stellmack, M. A.
Stelmachowicz, P. G.
Stepanishen, P. R.
Stergiopoulos, S.
Stevens, K. N.
Stewart, K.
Stinson, M. R.
Stoel-Gammon, C.
Stokes, A. P.
Stoll, R. D.
Stone, M. L.
Story, B. H.
Stotts, S. A.

Stove, L.
Strange, W.
Strasberg, M.
Strawderman, W. A.
Strelioff, D.
Strickland, E.
Strifors, H. C.
Strik, H.
Strong, W. J.
Strybel, T. Z.
Studdert-Kennedy, M. G.
Studebaker, G. A.
Sullivan, E.
Sulter, A. M.
Summerfield, A. Q.
Summers, I.
Summers, V.
Sundberg, J.
Sung, C-C.
Sussman, H.
Sussman, J.
Sutherland, L. C.
Suthers, R. A.
Svirsky, M. A.
Swanson, D. C.
Swift, G.
Swingler, D. N.
Syrdal, A. K.
Szabo, T. L.
Szwerc, R.
Szymko, Y. M.
Tabrikian, J.
Takahashi, D.
Talmadge, C. L.
Talmant, M.
Tan, J.
Tanaka, N.
Tang, D.
Tang, X. M.
Tappert, F.
Tavolga, W. N.
Teghtsoonian, M.
Telschow, K.
Temkin, S.
Ternstrom, S.
Tew, R. H.
Tewary,V. K.
Thibodeau, L.
Thijssen, J. M.
Thomas, D. R.
Thompson, G.
Thompson, L. L.
Thompson, P. O.
Thompson, S. C.
Thompson, W.
Thomson, D. J.
Thorne, P. D.
Thornton, A. R.
Thorsos, E. I.
Thuraisingham, R. A.
Tichy, J.
Tiersten, H. F.
Tindle, C. T.
Ting, R. Y.
Tittman, B. R.
Titze, I. R.
Tjøtta, J.
Tjøtta, S.
Tobey, E.
Tobocman, W.
Todd, N.
Tohkura, Y.
Tohkura, Y-i.
Toksöz, M. N.
Tolstoy, A. I.
Tolstoy, I.
Tong, J.
Toole, F. E.
Torston, D.
Trahey, G. E.
Trahiotis, C.
Traunmüller, H.
Trehub, S.
Trethewey, M.
Trevorrow, M. V.
Trinh, E.
Trivett, D. A.
Trudeau, M.
Tubis, A.
Tucci, D. L.
Tugnait, J. K.
Tuller, B.
Turner, C.
Turner, J.
Tyack, P.
Tye-Murray, N.
Tyler, R.
Uberall, H.
Ungar, E. E.
Uscinski, B. J.
Uziel, A.
Vaccaro, R.
Vagle, S.
Van den Abeele, K.
van den Honert, C.
Van Dijk, P.
Van Heuven, V.
Van Hoesel, R. J. M.
van Hoof, H. A. J. M.
van Lieshout, P. H. H. H.
van Santen, J.
van Son, R. J. J. H.
Van Tasell, D.
van Wieringen, A.
Varadan, V. V.
Varade, A.
Vardoulakis, I.
Vatikiotis-Bateson, E. A.
Veksler, N. D.
Verheij, J. W.
Verrillo, R. T.
Vidmar, P.
Vieira, M. N.
Viemeister, N.
Vilkman, E.
Villchur, E.
Visscher, W. M.
Viswanathan, E.
Voigt, H. F.
Vorlander, M.
Voronovich, A. G.
Vray, D.
Waag, R. C.
Wakefield, G.
Walden, B. E.
Walker, W. F.
Walley, A. C.
Walsh, E. J.
Wandelt, R.
Wang, B-T.
Wang, C-Y.
Wang, L.
Wang, P. P.
Wang, W. D.
Wapenaar, C. P. A.
Ward, W. D.
Warnock, A. C. C.
Warren, R.
Warren, R. M.
Washburn, K.
Watkins, A.
Watson, B. C.
Watson, C.
Watts, G. R.
Wear, K. A.
Weaver, R.
Weaver, R. L.
Webb, J. C.
Weber, D.
Webster, D. B.
Weinberg, H.
Weinreich, G.
Weinstein, B.
Weisenberger, J.
Weismer, G. G.
Wenzel, A. R.
Wenzel, E. M.
Werner, L. A.
Westbury, J. R.
Weston, D. E.
Westwood, E. K.
Whalen, D.
White, B.
White, M. J.
White, P.
Wiederhold, M. L.
Wightman, F. L.
Wild, M.
Williams, E. G.
Williams, K. L.
Wilson, B.
Wilson, D. K.
Wilson, J. H.
Wilson, O. B., Jr.
Wit, H. P.
Wong, D.
Wong, G. S. K.
Wong, K. M.
Woodhouse, J.
Worcester, P.
Wright, B. A.
Wu, J-r.
Wu, J.
Wu, S. F.
Wu, T-T.
Wu, T-W.
Wurmser, D.
Xiang, N.
Xie, Y.
Yamada, K.
Yamada, R. A.
Yamamoto, T.
Yang, T. C.
Yang, T.-Q.
Yang, Y.
Yates, G. K.
Ye, Z.
Yevick, D.
Yoon, S. W.
Yoshikawa, S.
Yost, W. A.
Yost, W. T.
Young, E. D.
Young, R. W.
Yun, K.
Zabolotskaya, E. A.
Zagzebski, J. A.
Zakarauskas, P.
Zarembowitch, A.
Zeng, F-G.
Zenner, H. P.
Zeroug, S.
Zhang, C.
Zhang, Tao
Zhou, G.
Zhou, J.-X.
Zhu, Q.
Zhu, R.-C.
Zingarelli, R. A.
Ziola, S.
Ziolkowski, R. W.
Ziomek, L. J.
Ziv, M.
Zolotov, E. M.
Zsiga, E.
Zuckerwar, A. J.
Zurek, P.
Zwislocki, J. J.
Zwolan, T.

Report of the Auditor

Published herewith is a condensed version of our auditor's report for calendar year ended 31 December 1996.

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, 1996 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, 1996 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.
12 May 1997
New York, NY

ACOUSTICAL SOCIETY OF AMERICA
STATEMENTS OF FINANCIAL POSITION
AS OF 31 DECEMBER 1996
(With Comparative Totals for 1995)

	1996	1995
Assets:		
Cash and cash equivalents	$ 681,225	$3,909,690
Accounts receivable	285,485	292,474
Marketable securities	3,978,955	35,230
Furniture, fixtures and equipment—net	112,526	105,484
Other assets	309,756	223,053
Total assets	$5,367,947	$4,565,931
Liabilities:		
Accounts payable and accrued expenses	$ 129,029	$ 249,102
Deferred revenue	1,048,038	887,657
Deferred rent liability	41,326	36,423
Total liabilities	$1,218,393	$1,173,182
Net assets:		
Unrestricted	$3,272,177	$2,638,424
Temporarily restricted	391,476	322,054
Permanently restricted	485,901	432,271
Total net assets	$4,149,554	$3,392,749
Total liabilities and net assets	$5,367,947	$4,565,931

ACOUSTICAL SOCIETY OF AMERICA
STATEMENTS OF ACTIVITY FOR THE YEAR ENDED 31 DECEMBER 1996
(With Comparative Totals for 1995)

	1996	1995
Unrestricted net assets:		
Revenue:		
Dues	$ 508,513	$ 461,103
Publishing	2,217,577	2,304,306
Standards	313,893	250,990
Meetings	305,231	176,603
Interest and dividends	133,759	268,271
Unrealized gain (loss)	252,935	140,873
Other	99,870	68,537
Realized gain (loss)	34,089	–
Total unrestricted revenue	$3,865,867	$3,670,683
Expenses:		
Publishing	$1,823,241	$1,731,188
Standards	411,513	380,511
Administrative and general	510,838	519,100
Meetings	313,000	224,495
Other expenses	177,563	129,792
Total expenses	$3,236,155	$2,985,086
Increase in net assets before other changes	$ 629,712	$ 685,597
Cumulative effect on prior years of adoption of SFAS No. 106	–	(79,100)
Increase in net assets	$ 629,712	$ 606,497
Net assets released from restrictions:		
Satisfaction of program restrictions	4,041	26,699
Increase in unrestricted net assets and reclassifications	$ 633,753	$ 633,196
Temporarily restricted net assets:		
Contributions	$ 20,301	$ 60,253
Investment income	21,213	32,382
Unrealized gain (loss)	31,949	22,376
Release of restrictions	(4,041)	(26,699)
Increase (decrease) in temporarily restricted net assets	$ 69,422	$ 88,312
Permanently restricted net assets:		
Investment income	$ 27,616	$ 43,463
Contributions		17,080
Unrealized gain (loss)	41,597	30,032
Expenses	(15,583)	(28,807)
Increase (decrease) in permanently restricted net assets	$ 53,630	$ 61,768
Increase in net assets	$ 756,805	$ 783,276
Net assets, beginning of year	$3,392,749	$2,609,473
Net assets, end of year	$4,149,554	$3,392,749

ACOUSTICAL SOCIETY OF AMERICA
STATEMENTS OF CASH FLOWS
FOR THE YEAR ENDED 31 DECEMBER 1996
(With Comparative Totals for 1995)

	Total All Funds	
	1996	**1995**
Operating Activities:		
Increase in net assets	$ 756,805	$ 783,276
Adjustments to reconcile net income to net cash provided by operating activities:		
Depreciation and amortization	23,973	21,004
Unrealized (gain) loss on marketable securities	(326,481)	(193,281)
Changes in operating assets and liabilities:		
(Increase) decrease in accounts receivable	6,989	(104,276)
Decrease in accrued interest receivable	–	34,361
(Increase) in other assets	(86,703)	(47,172)
Increase (decrease) in accounts payable and accrued expenses	(120,073)	151,207
Increase in deferred rent liability	4,903	4,903
Increase (decrease) in deferred revenue	160,381	(146,790)
Net cash flows provided by operating activities	$ 419,794	$ 503,232
Investing Activities		
Purchase of furniture, fixtures, equipment and leasehold improvements	($ 31,014)	($ 33,479)
Proceeds from sale of securities	2,097,286	2,177,095
Purchase of securities	(5,714,531)	–
Net cash (used in) provided by investing activities	($3,648,259)	$2,143,616
Increase (decrease) in cash and cash equivalents	($3,228,465)	$2,646,848
Cash and cash equivalents, beginning of year	3,909,690	1,262,842
Cash and cash equivalents, end of year	$ 681,225	$3,909,690

NOISE-CON 98 to be held in Ypsilanti, Michigan

NOISE-CON 98, the 1998 National Conference on Noise Control Engineering, will be held in Ypsilanti, Michigan on 5–8 April 1998. Ypsilanti is 15 miles west of the Detroit Metro Airport and seven miles east of Ann Arbor. Registration for the meeting will begin on Sunday, 5 April, and the technical sessions will begin on Monday morning, 6 April.

A major exhibition of instruments, software, materials, devices, and services for noise control will be held. It will open with a reception in the exhibition area on Monday evening, 6 April, and close on 8 April.

The conference is being organized and sponsored by the Institute of Noise Control Engineering (INCE) in cooperation with the Society of Automotive Engineers (SAE).

Robert J. Bernhard of Purdue University will be the General Chairman for the Conference. J. Stuart Bolton of Purdue and Paul Donavan of General Motors will serve as the Technical Program Chairs.

The theme of NOISE-CON 98 will be *Transporting Noise Control to the 21st Century: Planning for a Quiet Future*. Although technical papers in all areas of noise control engineering are welcome, papers on transportation vehicle noise sources, noise control materials, and noise control production are especially welcome.

NOISE-CON 98 will be followed by a SOUND QUALITY SYMPOSIUM on 9 April. Patricia Davies of Purdue University and Gordon Ebbitt of Lear Industries will be the General Chairs. Papers in all areas of sound quality are welcome. Topics to be covered include automotive sound quality, metrics for complex sounds, sound quality of motors, appliances, and components, and HVAC sound quality.

Deadline for receipt of abstracts for both meetings is 16 September 1997, and deadline for receipt of papers to be included in the proceedings of both the conference and the symposium is 16 January 1998.

The Announcement and Call for Papers is available from the Conference and Symposium Secretariat, Noise Control Foundation, P.O. Box 2469 Arlington Branch, Poughkeepsie, NY 12603. Telephone: (914) 462-4006; FAX: (914) 463-0201. E-mail: noisecon98@aol.com. The Announcement and Call for Papers can be found on the Internet at: http://users.aol.com/noisecon98/nc98_cfp.html

Regional Chapter News

Delaware Valley: The March meeting was held at the David Sarnoff Research Center in Princeton, NJ on Thursday, 20 March 1997 jointly with the ACM–IEEE Chapters. The program was about software for making phone calls via the Internet. This meeting was again arranged by DVASA Vice-President Rebecca Mercuri, and had surprises such as a one-week delay, during which the speaker, Dr. Ramalho left the company. However, Voxware sent an excellent substitute, Dr. Raymond Chen, Director of Research. He showed some superb computer graphics on the big screen at DSRC, and demonstrated some voice recordings processed by the software. Voices can be completely disguised (change your gender without changing your body). We learned that a Pentium 133 MHz computer is needed to run Voxware properly.

JON R. SANK

Open Position

Transportation—Related Acoustics Engineer or Physicist: The Acoustics Facility of the Volpe National Transportation Systems Center is seeking a senior-level professional to perform acoustic/noise studies related to transportation. These studies involve the modeling and measurement of noise emanating from aircraft and ground vehicles. Familiarity with noise models such as the Federal Highway Administration's STAMINA model and the Federal Aviation Administration's INM model are useful. Applicant should be cognizant of modern noise measurement and analysis instrumentation and methods. A degree, preferably in engineering or physics, from an accredited college or university is required. Salary ranges from $55,068 to $71,587. Send resumes to: Alfrida Coombs, DTS-841, Volpe Center, 55 Broadway, Cambridge, MA 02142. http://www.volpe.dot.gov The U.S. Government is an Equal Opportunity Employer.

USA Meetings Calendar

Listed below is a summary of meetings related to acoustics to be held in the U.S. in the near future. The month/year notation refers to the issue in which a complete meeting announcement appeared.

1997

21–27 Aug.	1997 Conference on Implantable Auditory Prostheses, Pacific Grove, CA [Alena Wilson, Conference Coordinator, House Ear Inst., 2100 W. 3rd St., Los. Angeles, CA 90057; Tel.: 213-353-7086; FAX: 213-413-0950; E-mail: alena@hei.org; WWW: http://www.rti.org/ciap97].
7–11 Sept.	American Academy of Otolaryngology—Head and Neck Surgery, San Francisco, CA [American Academy of Otolaryngology—Head and Neck Surgery, One Prince St., Alexandria, VA 22314; Tel.: 703-836-4444; FAX: 703-683-5100].
19–20 Sept.	Fifth Annual Conference on Management of the Tinnitus Patient, Iowa City, IA [Richard Tyler, Univ. of Iowa, Dept. of Otolaryngology—Head & Neck Surgery, 200 Hawkins Dr. C21-3GH, Iowa City, IA 52242-1078, Tel.: 319-356-2471; FAX: 319-353-7639; E-mail: rich-tyler@uiowa.edu].
22–24 Sept.	Second Biennial Hearing Aid Research and Development Conference, Bethesda, MD [National Institute of Deafness and Other Communication Disorders, 301-970-3844; FAX: 301-907-9666; E-mail: hearingaid@tascon.com]. Deadline for abstracts is 15 March.
1–5 Dec.	134th meeting of the Acoustical Society of America, San Diego, CA [ASA, 500 Sunnyside Blvd., Woodbury, NY 11797, Tel.: 516-576-2360; FAX: 516-576-2377; E-mail: asa@aip.org, WWW: http://asa.aip.org].

1998

9–13 Feb.	1998 Ocean Sciences Meeting, San Diego, CA [American Geophysical Union, 2000 Florida Ave., N.W., Washington, DC 20009, Tel.: 202-462-6900; FAX: 202-328-0566; WWW: http://www.agu.org].
5–8 April	NOISE-CON 98, Ypsilanti, MI [Noise Control Foundation, P.O. Box 2469, Arlington Branch, Poughkeepsie, NY 12603, Tel.: 914-462-4006; FAX: 914-463-0201; E-mail: noisecon98@aol.com, WWW: users.aol.com/noisecon98/nc98_cfp.html]. *8/97*
22–26 June	135th meeting of the Acoustical Society of America/16th International Congress on Acoustics, Seattle, WA [ASA, 500 Sunnyside Blvd., Woodbury, NY 11797, Tel.: 516-576-2360; FAX: 516-576-2377; E-mail: asa@aip.org, WWW: http://asa.aip.org].
7–12 July	Vienna and the Clarinet, Ohio State Univ., Columbus, OH [Keith Koons, Music Dept., Univ. of Central Florida, P.O. Box 161354, Orlando, FL 32816-1354, Tel.: 407-823-5116; E-mail: kkons@pegasus.cc.ucf.edu].
13–17 Sept.	American Academy of Otolaryngology—Head and Neck Surgery, San Francisco, CA [American Academy of Otolaryngology—Head and Neck Surgery, One Prince St., Alexandria, VA 22314, Tel.: 703-836-4444; FAX: 703-683-5100].
12–16 Oct.	136th meeting of the Acoustical Society of America, Norfolk, VA [ASA, 500 Sunnyside Blvd., Woodbury, NY 11797, Tel.: 516-576-2360; FAX: 516-576-2377; E-mail: asa@aip.org, WWW: http://asa.aip.org].

ACOUSTICAL NEWS—INTERNATIONAL

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

Unification of FASE and EAA

About 25 years ago a number of European Acoustical Societies or Associations formed the Federation of Acoustical Societies of Europe (FASE). FASE held joint symposia and other meetings which were usually organized by one or more of the member societies. The federation had officers but no individual members other than more than 20 European acoustical societies and some societies of audiology and a few commissions or groups of national academies of science.

With the advent of EAA, the European Acoustics Association, and EEAA, the East European Acoustics Association, some societies who were members of two of the three big groups suggested that simplification should be achieved in a manner satisfactory to all concerned. After many months of negotiations between FASE and EAA, the two organizations have merged and FASE has officially ceased to exist as of 1 January 1997. The former officers of FASE will serve out their terms as consultants to the Board of EAA in a purely advisory fashion without any legal rights within EAA.

The status of the still existing formal agreement of cooperation between FASE and EEAA has not been resolved yet. One of the difficulties is that EEAA had a special standing as an Associate Society of FASE and thus, not being a full member of FASE, could not be included in the merger.

A history of FASE will be prepared by Paul François (former Secretary of FASE) and R. C. Chivers (last President) and will be published by EAA in its journal ''Acta Acustica united with Acustica.''

International Conference on Shallow-Water Acoustics (SWAC '97)—Beijing, April 1997

The International Conference on Shallow-Water Acoustics (SWAC '97) was held at the Grand Hotel, Beijing, China on 21–25 April 1997. The conference was organized by the Chinese Academy of Sciences (Institute of Acoustics and National Laboratory of Acoustics), Georgia Institute of Technology (USA), and the Naval Postgraduate School (USA) and sponsored by the Office of Naval Research [Ocean Acoustics Program (USA)], Chinese Academy of Sciences, National Science Foundation of China, and the China State Shipbuilding Corporation. The honorary chairman was Prof. Dezhao Wang from the Institute of Acoustics, Chinese Academy of Sciences. The conference chairman was Prof. Renhe Zhang from the National Laboratory of Acoustics, Chinese Academy of Sciences.

Over the 5 days of the conference a broad range of topics on shallow-water acoustics were covered which focused on the unique environments associated with all aspects of the shallow-water acoustics. Nonparallel sessions were used to assure conference participants had access to each paper. After the opening ceremony, which included presentations by the Chinese Academy of Sciences and the US Office of Naval Research, the conference was presented in a series of eight sequential sessions: Chinese—USA Cooperative Research, Theoretical and Computation Acoustics, Sound Propagation and Seafloor Characterization, Effects of Water Column Variability, Reverberation and Bottom Scattering, High-Frequency Acoustics/Scattering from Surfaces and Bubbles/Ambient Noise, Acoustic Tomography/Signal Processing and Instruments, and Signal Processing. Each session started with invited papers which set the framework for the contributed papers that followed.

The over 100 attendees were provided with an exceptional social and cultural program orchestrated by the conference host. The program started with an opening reception which included an enlightening presentation on the history of Beijing. Evenings included working dinners with one evening including a performance of the Beijing Opera. In-depth tours of the Institute of Acoustics/National Laboratory of Acoustics were provided which focused on ongoing theoretical and experimental research. The evening of the final conference day included a closing ceremony and banquet at the Summer Palace. Daily tours to sites ranging from the Temple of Heaven to the Summer Palace were part of the social program for accompanying persons.

The conference was a success due to the depth of the technical program and the tireless efforts of the Chinese Academy of Sciences. The conference proceedings will be published and distributed in approximately 4 to 5 months.

THOMAS H. NEIGHBORS
Science Applications International Corporation
McLean, VA 22102

Papers published in JASJ(E)

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

The May issue of JASJ(E), Vol. 18, No. 3 (1997) contains the following papers:

Z. Maekawa ''Environmental acoustics update''
Robert T. Beyer ''Nonlinear threads in the coat of acoustics''
K. Ueda, T. Tanaka, and M. Orimoto ''Estimation of auditory source width for two adjacent 1/3 octave band noises with different band levels''
H. Imaizumi, M. Kinoshita, S. Kunimatsu, and T. Isei ''Sound propagation and speech hearing in a curved reverberant tunnel''

The July issue of JASJ(E), Vol. 18, No. 4 (1997) contains the following papers:

M. Akiho, M. Haseyama, and H. Kitajima ''ARMA direct filter designing method to reconstruct FIR digital filters based upon genetic algorithm and simulated annealing''
M. Ebata, H. Myazono, S. Suzuki, T. Usagawa, and Bertram Scharf ''Auditory detection of multiple targets''
H. Morikawa ''Analysis, synthesis, and perception of Japanese voiceless stop consonants based on a pole-zero model''
Y. Mitani, N. Nakasako, and M. Ohta ''A precise estimation method of probability distribution based on roughly observed data with quantized level and its application to acoustic measurement''

International Meetings Calendar

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

August 1997

16–19	**Fechner Day '97: Annual meeting of the International Society for Psychophysics**, Poznan. *6/97*
18–22	**3rd EUROMECH Solid Mechanics Conference**, Stockholm. *10/96*
19–22	**International Symposium on Musical Acoustics**, Edinburgh. *8/96*
21–23	**ACTIVE 97 Inter-Noise Satellite Symposium**, Budapest. *6/96*
24–27	**1997 World Congress on Ultrasonics**, Yokohama. *4/96*
25–27	**Inter-Noise 97**, Budapest. *2/96*
29–31	**Pan-European Voice Conference**, Regensburg. *6/97*

September 1997

1–4	**Modal Analysis Conference** (IMAC-XV Japan), Tokyo. *10/96*
9–12	**31st International Acoustical Conference ''Acoustics—High Tatra 97,''** High Tatra. *10/96*
10–12	**Biennial Conference New Zealand Acoustical Society**, Christchurch. *4/96, 12/96*

 0001-4966/97/102(2)/674/2/$10.00

10–12 **Biomechanics of Hearing**, Stuttgart. *10/96*
10–12 *__British Society of Audiology Annual Conference__, Cardiff, UK. (BSA, 80 Brighton Road, Reading RG6 1PS, UK; Fax: +44 118 936 1915)
15–18 **3rd EUROMECH Fluid Mechanics Conference**, Göttingen. *10/96*
16–19 **44th Open Seminar on Acoustics**, Jastrzebia Gora. *6/97*
17–19 *__Autumn Meeting of the Acoustical Society of Japan__, Hokkaido, Japan. (ASJ, Ikeda Bldg., 2-7-7, Yoyogi, Shibuya-ku, Tokyo, 151 Japan; Fax: +81 3 3379-1456)
18–19 **4th Mexican Congress on Acoustics**, Guanajuato. *4/96*
18–20 **Intonation: Theory, Models and Applications**, Athens. *2/97*
22–25 **European Conference on Speech Communication and Technology**, Patras. *2/96*
23–26 **Fluid-structure Interaction in Acoustics**, Delft. *10/96*
27–28 **Audio-Visual Speech Processing**, Rhodes. *6/97*

October 1997

7–10 **1997 IEEE Ultrasonics Symposium**, Toronto. *2/96*
8–10 **Acoustics Week in Canada 1997**, Windsor. *12/96*
23–26 **Reproduced Sound 13**, Windermere. *2/97*
30–31 **Swiss Acoustical Society Meeting**, Bern. *6/97*

November 1997

19–21 **WESTPRAC'97**, Hong Kong. *10/96*
27–30 **IOA Autumn Conference: Environmental Noise**, Windermere. *2/97, 6/97*

December 1997

15–18 **5th International Congress on Sound and Vibration**, Adelaide. *10/96*

February 1998

2–6 **Ultrasonic Technological Processes-98**, Moscow. *6/97*

March 1998

23–27 **DAGA 98** (German Acoustical Society Meeting), Zürich. *8/96*

May 1998

18–22 *__7th Spring School on Acoustooptics and Applications__, Gdańsk, Poland. (B. Linde, Institute of Experimental Physics, University of Gdańsk, ul. Wita Stwosza 57, 80-952 Gdańsk, Poland; Fax: +48 58 41 31 75; e-mail: school@univ.gda.pl)
25–27 **Noise and Planning 98**, Naples. *2/97*

June 1998

8–10 **EAA/EEAA Symposium ''Transport Noise and Vibration,''** Tallinn. *10/96*
9–12 **8th International Conference on Hand-Arm Vibration**, Umea. *6/97*
20–28 **Joint Meeting of the 16th International Congress on Acoustics and 135th Meeting of the Acoustical Society of America**, Seattle. *6/97*

November 1998

16–18 **Inter-Noise 98**, Christchurch. *4/96*
23–27 **ICBEN 98: Biological Effects of Noise**, Sydney. *12/96*
30–4 **5th International Conference on Spoken Language Processing**, Sydney. *6/97*

March 1999

15–19 **Joint Meeting of EAA Forum Acusticum and 137th Meeting of the Acoustical Society of America**, Berlin. *6/97*

REVIEWS OF ACOUSTICAL PATENTS

Daniel W. Martin
7349 Clough Pike, Cincinnati, Ohio 45244

The purpose of these acoustical patent reviews is to provide enough information for a Journal reader to decide whether to seek more information from the patent itself. Any opinions expressed here are those of reviewers as individuals and are not legal opinions. Printed copies of United States Patents may be ordered at $3.00 each from the Commissioner of Patents and Trademarks, Washington, DC 20231. [S0001-4966(97)03308-0]

Reviewers for this issue:

MAHLON D. BURKHARD, *31 Cloverdale Heights, Charles Town, West Virginia 25414*
HARVEY H. HUBBARD, *325 Charleston Way, Newport News, Virginia 23606*
SAMUEL F. LYBARGER, *101 Oakwood Road, McMurray, Pennsylvania 15317*
D. LLOYD RICE, *11222 Flatiron Drive, Lafayette, Colorado 80026*
CARL J. ROSENBERG, *Acentech Incorporated, 33 Moulton Street, Cambridge, Massachusetts 02138-1118*
ERIC E. UNGAR, *Acentech Incorporated, 33 Moulton Street, Cambridge, Massachusetts 02138-1118*
ROBERT C. WAAG, *University of Rochester Medical Center, 601 Elmwood Avenue, Rochester, New York 14642*

5,457,752

43.30.Yj DRIVE SYSTEM FOR ACOUSTIC DEVICES

Göran Engdahl *et al.*, assignors to ABB Atom AB
10 October 1995 (Class 381/190); filed in Sweden 29 August 1991

The drive system contains a number of drive devices placed in parallel and symmetrically between pressure beams which are included in these devices. The drive devices comprise a fixture frame inside of which there are two drive units with an intermediate mechanical prestress device. By stacking a number of driving elements on top of each other, drive units with different lengths may be constructed.—MDB

5,446,333

43.35.Zc ULTRASONIC TRANSDUCERS

Yoshihiko Ishida and Makoto Tani, assignors to NGK Insulators, Ltd.
29 August 1995 (Class 310/334); filed in the United Kingdom 21 September 1992

Claim one reads in part: ''An ultrasonic transducer unit for use in pulse-echo ultrasonic investigation at an operating frequency, comprising at least a piezoelectric ultrasonic wave transmitting element, an acoustoelectric ultrasonic wave receiving element consisting essentially of a ZnO single crystal, and electrodes therefor bonded together in an integrated multi-layer structure in which said transmitting element and said receiving element are superimposed in the direction of propagation... .''—MDB

5,447,069

43.35.Yb APPARATUS AND METHOD FOR ULTRASONICALLY MEASURING THE POISSON'S RATIO OF THIN LAYERS

Eric C. Johnson *et al.*, assignors to The Aerospace Corporation
5 September 1995 (Class 73/602); filed 29 June 1993

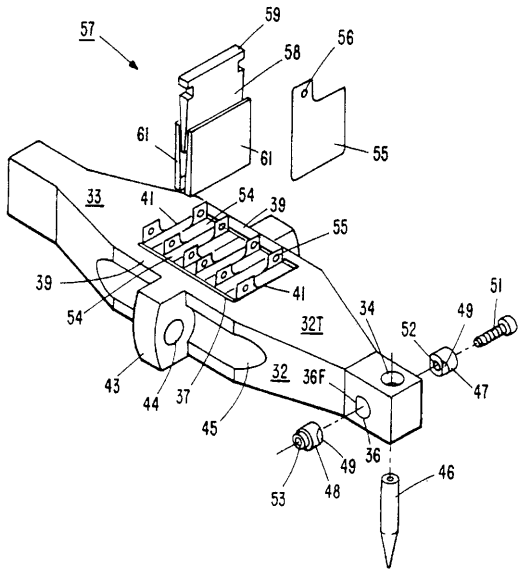
An ultrasonic resonance technique which permits measurement of the Poisson's ratio of thin adhesive material specimens is described. The acoustic shear and longitudinal velocity are also determined with the system. A fluid medium is used to couple sound into the specimen.—MDB

5,469,011

43.35.Zc UNIBODY ULTRASONIC TRANSDUCER

Ali R. Safabakhsh, assignor to Kulicke & Soffa Investments
21 November 1995 (Class 310/325); filed 6 December 1993

A rectangular aperture **37** is cut into the center of a single piece of metal **33**. A stack of piezoelectric crystals is mounted in the aperture so as to



5,465,468

43.35.Yb METHOD OF MAKING AN ELECTROMECHANICAL TRANSDUCER DEVICE

Ronald R. Manna, assignor to Misonix, Incorporated
14 November 1995 (Class 29/25.35); filed 28 September 1993

A liquid cooled transducer construction is shown. Cooling liquid passes through the center of the transducer. Piezoelectric ceramic driving elements are concentric to the main flow path. O-rings seal the central flow path from the volume occupied by the driving elements.—MDB

 0001-4966/97/102(2)/676/13/$10.00

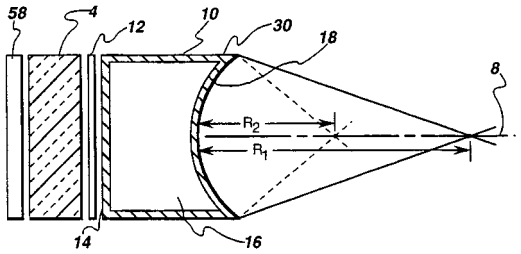
contact only the ends of the aperture. Bonding tool **46** is firmly anchored to the transducer.—MDB

5,477,736

43.35.Zc ULTRASONIC TRANSDUCER WITH LENS HAVING ELECTRORHEOLOGICAL FLUID THEREIN FOR DYNAMICALLY FOCUSING AND STEERING ULTRASOUND ENERGY

Peter W. Lorraine, assignor to General Electric Company
26 December 1995 (Class 73/642); filed 14 March 1994

An electrorheological fluid has voltage-dependent acoustical properties which can, for example, alter the velocity of sound, attenuation or the nonlinearity in the fluid. In this case the electrorheological fluid consists of dielectric particles floating in an insulating fluid having voltage-dependent



flow properties. By placing a lens **10** containing the fluid **16** in front of the transducer element **4**, the focal point of the lens may be altered by application of a voltage to the lens.—MDB

5,479,825

43.35.Zc TRANSDUCERS FOR TRANSMITTING AND/OR RECEIVING ULTRASONIC ENERGY

Paul Williams and Bradley M. Pankonin, assignors to ABB Industrial Systems
2 January 1996 (Class 73/644); filed 28 January 1993

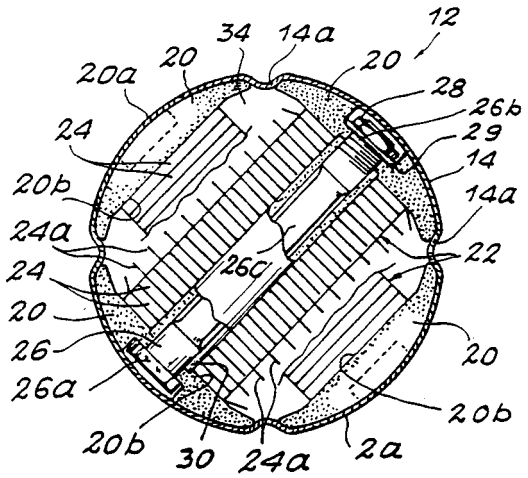
An application of the transducers disclosed in this patent is the ultrasonic measurement of physical properties of sheets of materials, e.g., webs of paper as they are being manufactured. The transducers are constructed using rectangular bars or circular discs of piezoelectric material.—MDB

5,477,101

43.38.Fx DOWNHOLE ACOUSTIC TRANSDUCER

Abderrhamane Ounadjela, assignor to Schlumberger Technology Corporation
19 December 1995 (Class 310/334); filed in France 6 November 1990

An acoustic transducer intended for use in a well, e.g., an oil well, to generate or detect acoustic waves. Depending on the application, the transducer may act as a seismic source or receiver. The figure shows in cross



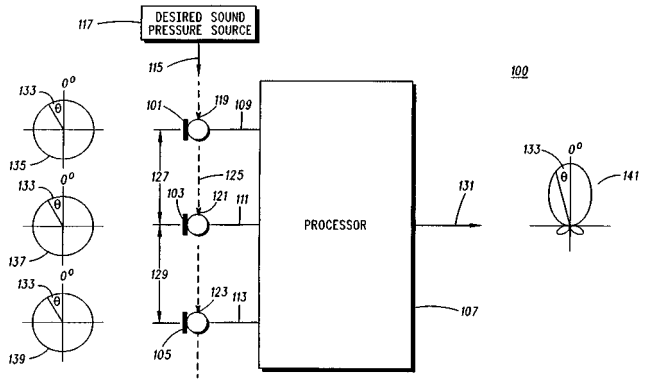
section one stack of piezoelectric elements **22** and a second stack **22** below it and at 90° orientation, these being two of a number of stacks, encased in a flexible and water tight tubular housing **2a**.—MDB

5,463,694

43.38.Hz GRADIENT DIRECTIONAL MICROPHONE SYSTEM AND METHOD THEREFOR

Wayne H. Bradley and Richard E. Werner, assignors to Motorola
31 October 1995 (Class 381/92); filed 1 November 1993

Output signals from no more than three similar microphones are processed to produce direction-sensitive pickup of a sound signal. The proces-



sor output is expected to be of higher-order gradient than the gradient of the respective microphones.—MDB

5,473,701

43.38.Hz ADAPTIVE MICROPHONE ARRAY

Juergen Cezanne and Gary W. Elko, assignors to AT&T Corporation
5 December 1995 (Class 381/92); filed 5 November 1993

The patent describes a technique for adaptively adjusting the directivity of a microphone array to reduce the sensitivity of the array to background noise. An example includes an array having a number of microphones. The directivity pattern of the array may be adjusted by varying one or more parameters so as to create desirable angular orientations of one or more directivity pattern nulls. Evaluation of parameters is performed under a constraint that the orientation of the nulls is restricted to a predetermined region termed the background.—MDB

5,455,869

43.38.Kb LAVALIER MICROPHONE ASSEMBLY

David Miscavige, assignor to Church of Scientology International
3 October 1995 (Class 381/169); filed 19 March 1990

A cage or open shell structure surrounds a microphone to reduce vibration noise pickup by the microphone due to rubbing against clothing while providing for attaching to the garment.—MDB

5,473,684

43.38.Kb NOISE-CANCELLING DIFFERENTIAL MICROPHONE ASSEMBLY

Charles S. Bartlett and Michael A. Zuniga, assignors to AT&T Corporation
5 December 1995 (Class 379/387); filed 21 April 1994

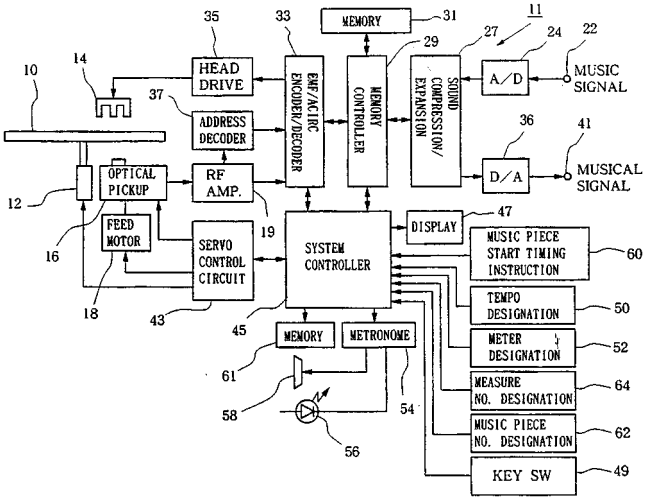
With emphasis on improved pickup of close sound sources while decreasing pickup of distant sounds, as in a telephone handset, for example, ''improved microphone performance is achieved by configuring a second-order derivative microphone assembly in such a way that radially divergent near-field input produces a microphone response proportional to a first-order spatial derivative of the acoustic pressure field.''—MDB

5,602,811

43.38.Md MUSICAL INFORMATION RECORDING AND REPRODUCING TECHNIQUE FOR USE WITH A RECORDING MEDIUM HAVING A UTOC AREA

Mikio Ogusu *et al.*, assignors to Yamaha Corporation
11 February 1997 (Class 369/47); filed in Japan 11 August 1994

This digital recording and reproducing system uses an optical Mini Disk for recording various parts in a musical performance in parallel independently for mixing and editing purposes and for selective quick access to any segment of the musical recording. ''There is generated a rhythm corresponding to a designated tempo and meter of a music piece to be recorded. Performance of a musical instrument is started at specific timing according to the generated rhythm, and the musical information resulting from the performance is recorded onto an MD (Mini Disk). The address on the MD readout when the performance of the music piece is started is retrieved as a start address of the music piece. After the recording, the retrieved start address is recorded onto the UTOC (User's-table-of-contents) area of the disk along with the designated tempo and meter. By reading out the start address, tempo and meter to calculate a measure number and start address of a desired measure on the basis of the read out information, display of a measure number being currently reproduced and quick access to the starting point of any desired measure can be achieved. In applications where plural channels are involved, part information indicative of performance parts assigned and recorded on the individual recording channels is also recorded onto the UTOC area.''—DWM



5,471,538

43.38.Tj MICROPHONE APPARATUS

Tooru Sasaki and Kaoru Gyotoku, assignors to Sony Corporation
28 November 1995 (Class 381/92); filed in Japan 8 May 1993

''A microphone apparatus having a first microphone for picking up a desired sound [coming from an arrival direction] and a second microphone with directionality in which sensitivity is low to the desired sound arrival direction. A sound signal from the second microphone is supplied to a subtracting circuit through an adaptive filter. The subtracting circuit subtracts an output signal of the adaptive filter from the sound signal coming from the first microphone. A circuit is provided to adjust the adaptive filter so that the output power of the subtracting circuit is minimized.''—MDB

5,604,891

43.40.Rj 3-D ACOUSTIC INFINITE ELEMENT BASED ON A PROLATE SPHEROIDAL MULTIPOLE EXPANSION

David S. Burnett and Richard L. Holford, assignors to Lucent Technologies, Incorporated
18 February 1997 (Class 395/500); filed 27 October 1994

This patent describes an improvement over conventional finite element techniques that are useful in the field of computational structural acoustics. ''A prolate spheroidal infinite element is used for the modeling of acoustic fields in exterior, fluid-filled domains surrounding a structure. This prolate infinite element is based on a multipole expansion that describes, to arbitrary accuracy, any scattered and/or radiated field exterior to a prolate spheroid of any eccentricity. The prolate infinite element is readily incorporated in any structural or acoustical finite element code. Structural acoustic modeling with the element is several orders of magnitude faster than modeling, at comparable accuracy, with the well-known and widely used boundary element method, at least for large-scale problems involving tens of thousands of degrees of freedom.'' For supporting documentation see J. Acoust. Soc. Am. **96**, 2798–2816 (1994).—DWM

5,604,893

43.40.Rj 3-D ACOUSTIC INFINITE ELEMENT BASED ON AN OBLATE SPHEROIDAL MULTIPOLE EXPANSION

David S. Burnett and Richard L. Holford, assignors to Lucent Technologies, Incorporated
18 February 1997 (Class 395/500); filed 18 May 1995

This patent is a companion to patent 5,604,891 reviewed above, sharing the same assignors, assignee, and many of the same figures. This patent uses an oblate, rather than prolate, spheroidal multipole expansion.—DWM

5,558,191

43.40.Tm TUNED MASS DAMPER

Ming-Lai Lai, assignor to Minnesota Mining and Manufacturing Company
24 September 1996 (Class 181/379); filed 9 March 1995

This patent pertains to dynamic absorbers, also called ''tuned dampers,'' that incorporate viscoelastic elements. In such a damped spring-mass system, which is to be attached to an item having vibrations needing to be suppressed in a narrow frequency range, the viscoelastic element may serve as both spring and damper, or primarily just as a damper. The patent illus-

trates a number of conceptual configurations of dynamic absorbers that operate torsionally or translationally in two or three directions.—EEU

5,558,497

43.40.Tm AIRFOIL VIBRATION DAMPING DEVICE

Robert J. Kraft and Robert J. McClelland, assignors to United Technologies Corporation
24 September 1996 (Class 416/96 A); filed 31 July 1995

The damper described in this patent consists of a reed-shaped element that is built into a rotor blade, and that is intended to damp vibrations of that blade as the result of friction between it and the blade. In one version the damper is pressed against an internal surface of the blade by a pressure difference across it that results from suitable venting of cavities in the rotor blade. Another version relies on centrifugal force to press the damper against a mating internal surface of the blade.—EEU

5,573,220

43.40.Tm ADJUSTABLE VIBRATION ABSORBING MACHINERY FOUNDATION MOUNT AND METHOD FOR TUNING THE SAME

Wayne H. Whittaker *et al.*, assignors to Unisorb, Incorporated
12 November 1996 (Class 248/638); filed 30 May 1995

This heavy-duty machinery mount is in the form of a rectangular platform that is supported on legs at its four corners. These legs consist of rectangular tubes that extend above the platform and that can be filled with various amounts of different types of resilient pads. Thus the bottom-most pads make contact with the floor that supports the entire system, and the stacks of resilient pads provide isolation. A vertical level-adjusting device consisting of a screw-actuated wedge arrangement is incorporated into the mounting system.—EEU

5,573,334

43.40.Tm HIGH DAMPING COMPOSITE JOINT FOR MECHANICAL VIBRATION AND ACOUSTIC ENERGY DISSIPATION

Roger M. Crane *et al.*, assignors to the United States of America
12 November 1996 (Class 403/179); filed 17 October 1994

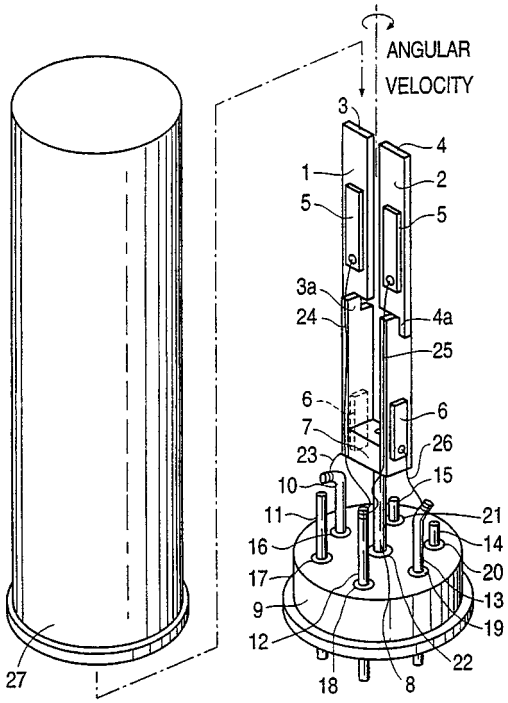
In a typical, nondissipative joint the end of one member (such as a tube) fits into a matching opening in the member to which it is to be joined, and an adhesive in the clearance space serves to connect the two. In a joint described by this patent the interface gap is longer than usual, and some of it is filled with a viscoelastic material. The dimensions and material properties are chosen so that as the axial load on the joint increases, and the adhesive becomes increasingly plastic, the load is increasingly taken by the viscoelastic material. Then vibrations superposed on this static loading are damped due to energy dissipation in the viscoelastic material.—EEU

5,447,066

43.40.Yq ANGULAR VELOCITY SENSOR HAVING A TUNING FORK CONSTRUCTION AND ITS METHOD OF MANUFACTURE

Jiro Terada *et al.*, assignors to Matsushita Electric Industrial Company
5 September 1995 (Class 73/504.16); filed 1 September 1993

Tuning fork vibrators **3** and **4** comprise metal base plates with 90° bends at **3a** and **4a**. Piezoelectric elements **5** and **6** are bonded to the base plates. An ac signal applied to elements **6** produces vibration of the base



plates, and Coriolis forces on the base plates due to rotation generate output signals through elements **5** that relate to rotation velocity.—MDB

5,447,067

43.40.Yq ACCELERATION SENSOR AND METHOD FOR MANUFACTURING SAME

Markus Biebl *et al.*, assignors to Siemens Aktiengesellschaft
5 September 1995 (Class 73/514.32); filed in Germany 30 March 1993

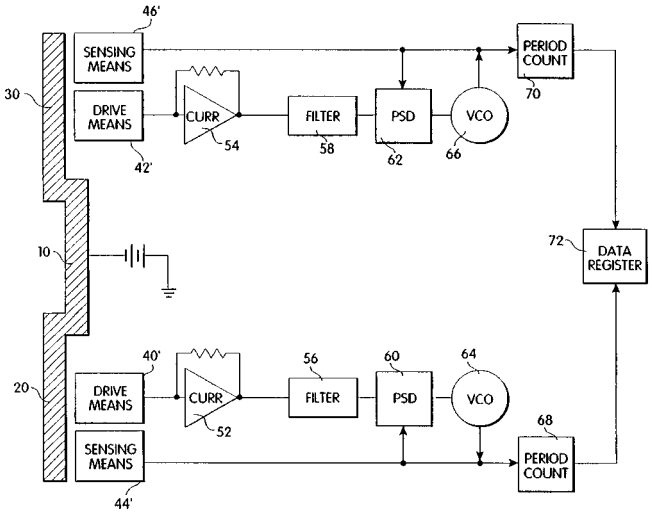
The acceleration sensor has a proof mass supported by flexible elements. It is fabricated by micromachining of doped monocrystalline silicon on an insulating substrate. Sensing of movement of the proof mass is by means of piezoresistors on the flexible elements.—MDB

5,461,918

43.40.Yq VIBRATING BEAM ACCELEROMETER

George Mozurkewich, assignor to Ford Motor Company
31 October 1995 (Class 73/514.26); filed 26 April 1993

The accelerometer comprises a driven cantilever beam and a sensor to detect resonant flexural vibrations of the beam. Acceleration along the major axis of the beam causes changes of resonance frequency. The magnitude of



the change in resonance frequency may be measured by a signal processing circuit whose output may be used to calculate the component of acceleration along the major axis of the beam.—MDB

5,456,109

43.40.Yq THICK FILM ROTATIONAL ACCELEROMETER HAVING TWO STRUCTURALLY INTEGRATED LINEAR ACCELERATION SENSORS

Frans P. Lautzenhiser *et al.*, assignors to Delco Electronics Corporation
10 October 1995 (Class 73/514.03); filed 29 March 1993

''A rotational accelerometer has a pair of coplanar, operationally independent linear acceleration sensors. Coplanarity is provided by a unitary substrate which forms the base structure of the operationally independent linear acceleration sensors.'' The sensors are piezoresistive electrical bridges and the substrate is alumina.—MDB

5,456,112

43.40.Yq HIGH ACCURACY LASER ACCELEROMETER

Joseph P. Ficalora, assignor to Allied Signal, Incorporated
10 October 1995 (Class 73/514.26); filed 21 December 1993

The laser accelerometer ''includes an acceleration sensing proof mass combined with a laser mirror, a laser or optically resonant cavity including a closed cavity, cathode, anode and a movable path length control mirror, a lens/prismatic readout to combine the two beams, a photo detector to convert the fringes to a sinusoidal electrical signal and associated electronics.''—MDB

5,477,729

43.40.Yq ACOUSTIC EMISSION TRANSDUCER

Claudio Cavalloni, assignor to K. K. Holding AG
26 December 1995 (Class 73/587); filed in Switzerland 2 September 1993

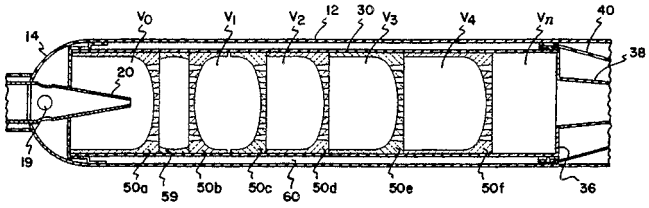
Intended for monitoring of machining processes by detecting the noises created thereby, a transducer is attached to a ''measuring object.'' In particular, an axially polarized piezoelectric element in the transducer is pressed onto the object so that a coupling diaphragm is preloaded with a reproducible force.—MDB

5,378,435

43.50.Gf SILENCER COMBINED WITH CATALYTIC CONVERTER FOR INTERNAL COMBUSTION ENGINES AND MODULAR DIAPHRAGM ELEMENTS FOR SAID SILENCER

Albino Gavoni, assignor to Gavoni B.G.M. Silenziatori Di Albino Gavoni
3 January 1995 (Class 422/177); filed in Italy 4 September 1991

This patent relates to the unique design of a combined exhaust silencer and catalytic converter for internal combustion engines. Provision is made for assembling a series of ceramic elements, such as inorganic refractory



oxides and carbides, in the axial direction for catalytic conversion and in the presence of a peripheral layer of rock wool or fiber glass for sound absorption.—HHH

5,410,927

43.50.Gf LOW NOISE PUNCH TOOL

Hitoshi Omata and Oriya Fujita, assignors to Amada Company
2 May 1995 (Class 83/139); filed 16 July 1993

This patent relates to a low-noise punch tool for mounting on a turret punch press. Conventional punch tools with shearing action cutters are fitted with urethane pads between the striker and the punch head between the punch driver and the punch guide, and between the stripper plate and the workpiece, in order to minimize noise radiation.—HHH

5,575,349

43.50.Gf SOUNDPROOF TYPE WATER COOLED ENGINE GENERATOR

Takatoshi Ikeda *et al.*, assignors to Denyo Kabushiki Kaisha
19 November 1996 (Class 180/68.1); filed in Japan 19 October 1993

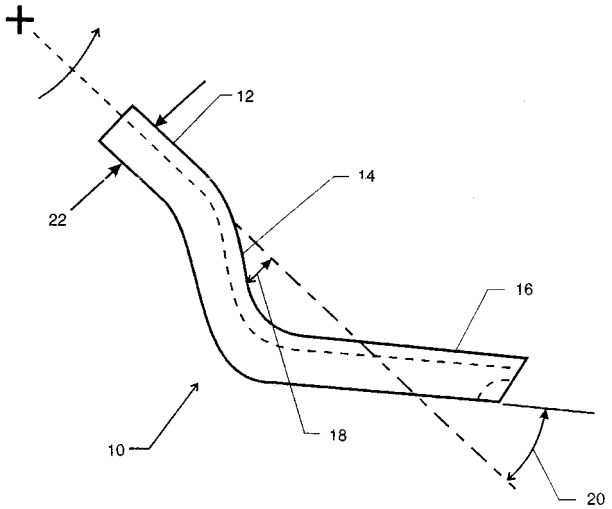
This patent applies to the noise control of a water-cooled engine generator. Cooling air is circulated through an enclosure around the engine and the generator, and exits through an exhaust muffler. A key noise reduction feature of the design is the location of the circulating fan remote from the radiator, in order to minimize the inflow irregularities to the fan.—HHH

5,584,661

43.50.Gf FORWARD SWEEP, LOW NOISE ROTOR BLADE

Thomas F. Brooks, assignor to the United States of America
17 December 1996 (Class 416/238); filed 6 July 1995

This patent relates to low-noise rotors for helicopters, particularly those for which impulsive blade-vortex interaction (BVI) noise may be significant. The preferred configuration incorporates about a 35° forward sweep over the outer 45% of the blade span. In the forward speed range for which



BVI noise dominates, the above geometry details can beneficially affect such noise generation factors as the strength and core size of the tip vortices at the times of interaction with a blade, and the interaction angles and Mach numbers.—HHH

5,597,985

43.50.Gf ACOUSTICALLY INACTIVE CORRUGATED STRUCTURE

Terrence A. Dear and Karl U. Ingard, assignors to E. I. DuPont de Nemours and Company
28 January 1997 (Class 181/227); filed 25 May 1995

The characteristic whistle of fluid flow through a corrugated structure (like a tube) is eliminated by varying the pitch of adjacent corrugations. The pitch of such a corrugated tube is defined as the distance between adjacent corrugations (peaks). The patent includes prescriptions on how the pitch should vary. Examples of corrugated tubes where the elimination of the whistle is most desired include medical applications such as performing aspiration in an operating room.—CJR

5,604,813

43.50.Ki INDUSTRIAL HEADSET

Robert D. Evans *et al.*, assignors to Noise Cancellation Technologies, Incorporated
18 February 1997 (Class 381/71); filed 2 May 1994

This industrial headset system is intended to provide both hearing protection through active noise reduction (ANR) and to provide two-way speech communication from audio or radio sources. The circuitry boosts the low-frequency response characteristic of the speech communication channel when the active noise cancellation is used simultaneously, to compensate for low-frequency loss which occurs as a result of active noise cancellation.—DWM

5,377,546

43.58.Bh MULTI-DEGREE LINEAR LINER IMPEDANCE TESTING DEVICE

Noe Arcas *et al.*, assignors to Grumman Aerospace Corporation
3 January 1995 (Class 73/589); filed 29 January 1992

There is claimed: ''1. Apparatus for evaluation of the acoustic impedance of structural members, comprising: spacer means for separating at least two of said structural members by a predetermined distance; alignment means for aligning an impedance tube with respect to said structural members; means for acoustically sealing said apparatus device such that acoustic signals generated within said impedance tube are directed onto said structural members.''—MDB

5,604,812

43.66.Ts PROGRAMMABLE HEARING AID WITH AUTOMATIC ADAPTION TO AUDITORY CONDITIONS

Wolfram Meyer, assignor to Siemens Audiologische Technik GmbH
18 February 1997 (Class 381/68.2); filed in European Patent Office 6 May 1994

A hearing aid is described that includes a first memory in which audiometric data are stored, a second memory in which hearing aid characteristics are stored, a third data memory in which algorithms are stored, and a signal analysis unit. It is stated that the characteristics of the amplifier/transmission circuit can be automatically determined from the edited audiometric data, prescribable algorithms, and the current ambient conditions.—SFL

5,606,620

43.66.Ts DEVICE FOR THE ADAPTATION OF PROGRAMMABLE HEARING AIDS

Oliver Weinfurtner, assignor to Siemens Audiologische Technik GmbH
25 February 1997 (Class 381/68.2); filed in European Patent Office 23 March 1994

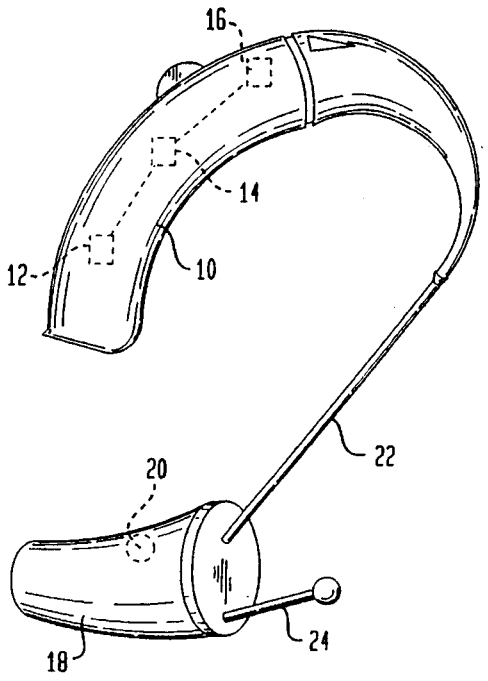
This device employs a fuzzy logic system for setting the characteristics of a programmable hearing aid according to the audiometric data of an individual. The device is a data processing unit with a fuzzy logic module having an arithmetic unit that processes the hearing impairment data and the characteristic data according to the principles of fuzzy logic. Circuitry and processing diagrams are shown.—SFL

5,606,621

43.66.Ts HYBRID BEHIND-THE-EAR AND COMPLETELY-IN-CANAL HEARING AID

James J. Reiter and Gordon Berkholcs, assignors to Siemens Hearing Instruments, Incorporated
25 February 1997 (Class 381/68.6); filed 14 June 1995

A hybrid BTE and CIC hearing aid has a BTE component which is worn behind the ear and a CIC component worn in the bony portion of the ear canal. The components are connected together with a wire cable. Acous-



tic feedback is said to be reduced, allowing gain to be increased. The occlusion effect is said to be reduced due to the placement of the CIC unit.—SFL

5,613,008

43.66.Ts HEARING AID

Raimund Martin, assignor to Siemens Audiologische Technik GmbH
18 March 1997 (Class 381/68.2); filed in Germany 29 January 1992

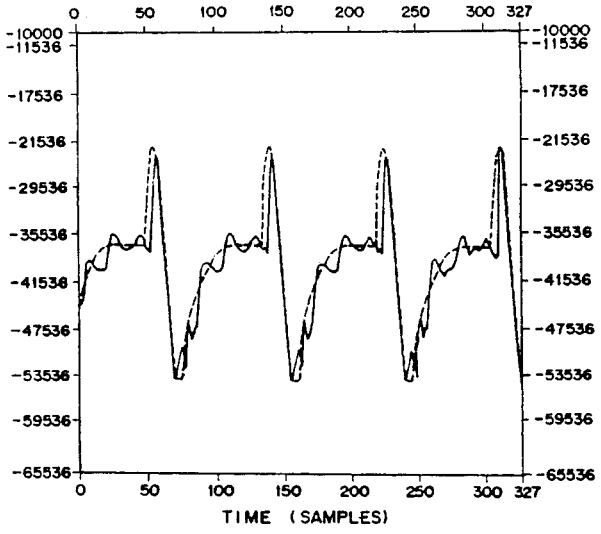
In a hearing aid, comprising a microphone, processing circuits for useful signals, an output stage, and transducer, the useful signals are sampled at a sampling frequency before the output stage, and converted into data words as well as being processed in discrete time fashion. The digital data words are capable of being converted into pulse-duration-modulated signals without being reconverted into analog signals. Circuitry is described.—SFL

5,577,160

43.72.Ar SPEECH ANALYSIS APPARATUS FOR EXTRACTING GLOTTAL SOURCE PARAMETERS AND FORMANT PARAMETERS

John-Paul Hosom and Mikio Yamaguchi, assignors to Sumitomo Electric Industries, Incorporated
19 November 1996 (Class 395/218); filed in Japan 24 June 1992

The patent describes various improvements to existing methods of speech analysis with the goal of combining these methods to achieve a highly accurate analysis procedure. Some of the highlights of the strategy are to extract formant one and glottal source parameters together, recognizing their mutual influence, to use only the closed glottis interval for extract-



ing formant frequencies and bandwidths, and to compensate for phase distortions introduced by analog recording techniques. In general, the strategy pursued is to filter out competing effects using preliminary estimates before each portion of the high quality analysis.—DLR

5,572,621

43.72.Dv SPEECH SIGNAL PROCESSING DEVICE WITH CONTINUOUS MONITORING OF SIGNAL-TO-NOISE RATIO

Rainer Martin, assignor to U.S. Philips Corporation
5 November 1996 (Class 395/2.36); filed in European Patent Office 21 September 1993

When operating a telephone or voice-activated device in a noisy environment, such as an automobile, an array of microphones may be arranged with adjustable delays from each such that the average of the delayed signals contains an improved signal-to-noise characteristic. A typical method for computing the individual delays depends on a good estimate of the signal-to-noise ratio (SNR) of each microphone signal. This device estimates the SNR of a signal by smoothing the power level, then detecting successive peaks and troughs in the smoothed power. The peak and trough measures are grouped by phoneme-length intervals such that they ideally represent signal and background noise, respectively.—DLR

5,572,623

43.72.Dv METHOD OF SPEECH DETECTION

Dominique Pastor, assignor to Sextant Avionique
5 November 1996 (Class 395/2.42); filed in France 21 October 1992

This patent discloses improvements to the well-known spectral subtraction method of noise reduction, consisting of new strategies for locating the beginning and end of voice activity. A FFT-based frame energy measure is used to locate a frame having a high certainty of being voiced speech. From that point backwards, a low-energy frame is located having energy near the long-term minimum and thus likely to be nonspeech. Various strategies at successively higher processing levels are then applied to examine the intermediate frames to locate the most likely moment of speech onset.—DLR

5,574,824

43.72.Dv ANALYSIS/SYNTHESIS-BASED MICROPHONE ARRAY SPEECH ENHANCER WITH VARIABLE SIGNAL DISTORTION

Raymond E. Slyh *et al.*, assignors to the United States of America
12 November 1996 (Class 395/2.35); filed 11 April 1994

The device consists of an array of microphones feeding a corresponding array of analysis filter banks. Signals from one or more sound sources may be selectively enhanced by choosing between two strategies for processing the filter band outputs. Both methods sum the outputs of the corresponding bands from all microphones, and apply a weighting pattern in the form of individual band gains before recombining the bands in a synthesis filter. The first weighting method is best for isolating a single source from uncorrelated background noise while the second is preferable when there are more sources than microphones.—DLR

5,577,161

43.72.Dv NOISE REDUCTION METHOD AND FILTER FOR IMPLEMENTING THE METHOD PARTICULARLY USEFUL IN TELEPHONE COMMUNICATIONS SYSTEMS

Clara S. Pelaez Ferrigno, assignor to Alcatel N. V.
19 November 1996 (Class 395/2.35); filed in Italy 20 September 1993

This noise reduction system improves the ratio of the speech signal to the uncorrelated background noise using a variation of the spectral subtraction method. Frames of the noisy speech input signal are windowed and Fourier transformed. Each frame is classified using the probability that speech is present given hypotheses that the present frame is similar to cur-

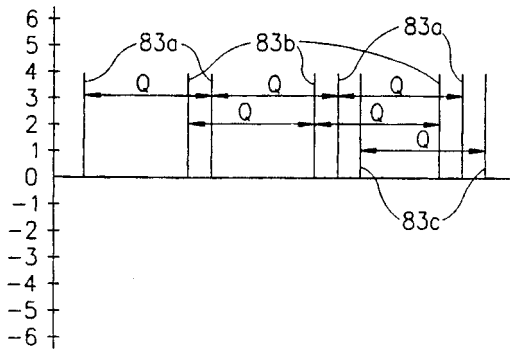
rent spectral shapes representing speech and noise, respectively. The current representative spectra are then selectively updated based on the frame classification. This provides a running estimate of the noise spectrum, which is then subtracted from the composite input signal.—DLR

5,568,588

43.72.Gy MULTI-PULSE ANALYSIS SPEECH PROCESSING SYSTEM AND METHOD

Leon Bialik and Felix Flomen, assignors to AudioCodes Limited
22 October 1996 (Class 395/2.32); filed 29 April 1994

This multipulse vocoder design includes two different versions, and a third which combines elements of the first two. In the first arrangement, a single excitation pulse is estimated from the initial linear prediction residual. A synthesized frame is then matched against the input, iterating over a few gain steps above and below the initial estimate, using a maximum likelihood criterion. Additional pulses are added, all having the same amplitude as the



first, but each time repeating a limited gain search, until a match acceptance level is reached. The second version duplicates a pattern of pulses at a time interval corresponding to the pitch period. The third combined version applies both strategies, using the best of the two methods for each frame.—DLR

5,570,453

43.72.Gy METHOD FOR GENERATING A SPECTRAL NOISE WEIGHTING FILTER FOR USE IN A SPEECH CODER

Ira A. Gerson *et al.*, assignors to Motorola, Incorporated
29 October 1996 (Class 395/2.28); filed 23 February 1993

Many current speech coder designs, and in particular the code-excited (CELP) vocoders, use a perceptual weighting filter in the trial synthesis loop, which has the effect of redistributing the parameter quantization noise such that it will be masked by the speech formants. A perceptual filter is generated for each short-term linear prediction vector. All codebook entries are then passed through both the LPC and the perceptual filters, resulting in a substantial computational load. This patent combines the two filters by passing an impulse through the filter pair, then using a Levinson recursion to find an approximation having lower order than the two original filters combined.—DLR

5,572,622

43.72.Gy REJECTED FRAME CONCEALMENT

Karl T. Wigren *et al.*, assignors to Telefonaktienbolaget LM Ericsson
5 November 1996 (Class 395/2.37); filed in Sweden 11 June 1993

Many current voice communications systems include a method, employed at the receiving terminal, of continuing ongoing sounds from the preceding frames to fill in for segments which have been lost due to excessive errors or late or missing packets. This patent describes two strategies for generating the fill-in sound depending on whether the last accepted frame was speech or background noise. For speech, the most recent frame energy and predictor coefficients are reused to extend the current speech sound, subject to certain timeouts. For nonspeech sounds, additional smoothing is applied to the energy and predictor parameters, producing a more averaged background sound.—DLR

5,572,681

43.72.Gy SPEECH CODEC AND A METHOD OF PROCESSING A SPEECH SIGNAL WITH SPEECH CODEC

Makio Nakamura and Akira Hioki, assignors to NEC Corporation
5 November 1996 (Class 395/2.21); filed in Japan 24 August 1992

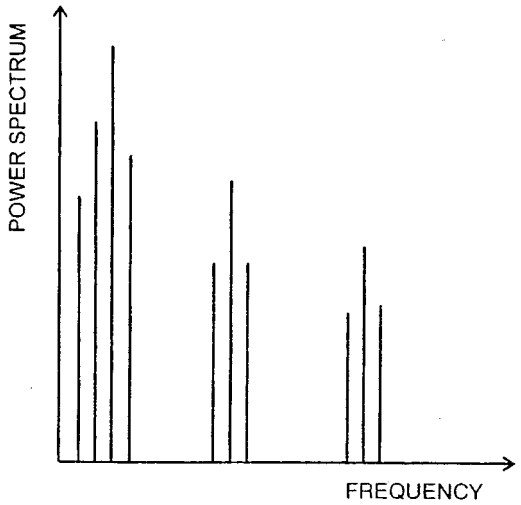
This patent describes the organization of a chip capable of encoding one stream of incoming speech samples (e.g., from an A/D converter) into a compressed format and simultaneously decoding a second stream of compressed speech into a stream of samples (e.g., for a D/A converter). Since the nature of the encoded compression format is not addressed here, the patent would perhaps be applicable to a range of techniques from logarithmic quantization up to, say, CELP vocoding.—DLR

5,574,823

43.72.Gy FREQUENCY SELECTIVE HARMONIC CODING

Hisham Hassanein *et al.*, assignors to Canada
12 November 1996 (Class 395/2.17); filed 23 June 1993

Harmonic coding consists of representing a voiced speech signal as a fundamental frequency value and a set of harmonic amplitude parameters covering some given frequency range. The vocoder presented here expands on that definition by a method of selecting a subset of the harmonic amplitudes to be transmitted. An alternative coding method presented is in fact a



quite different approach consisting of FFT power spectra which are divided into specific bands. The frequency positions of the selected bands are vector quantized with a band-position codebook. The patent concentrates primarily on describing the second coding method.—DLR

5,577,159

43.72.Gy TIME-FREQUENCY INTERPOLATION WITH APPLICATION TO LOW RATE SPEECH CODING

Yair Shoham, assignor to AT&T Corporation
19 November 1996 (Class 395/2.15); filed 9 October 1992

Early devices for representing speech in a time and frequency plane included the Haskins pattern playback and the widely used Sonograph. Recently, much effort has gone into mathematical representations that simul-

taneously take into account the time and frequency properties of the signal. One result of this is the development of more efficient coding techniques. Described here is a speech coder which uses the joint time-frequency properties to construct a low bitrate interpolated model of the speech spectral sequence. The result is a coder said to achieve quality-to-bitrate ratios roughly twice those of CELP coding.—DLR

5,574,825

43.72.Kb LINEAR PREDICTION COEFFICIENT GENERATION DURING FRAME ERASURE OR PACKET LOSS

Juin-Hwey Chen and Craig R. Watkins, assignors to Lucent Technologies, Incorporated
12 November 1996 (Class 395/2.37); filed 14 March 1994

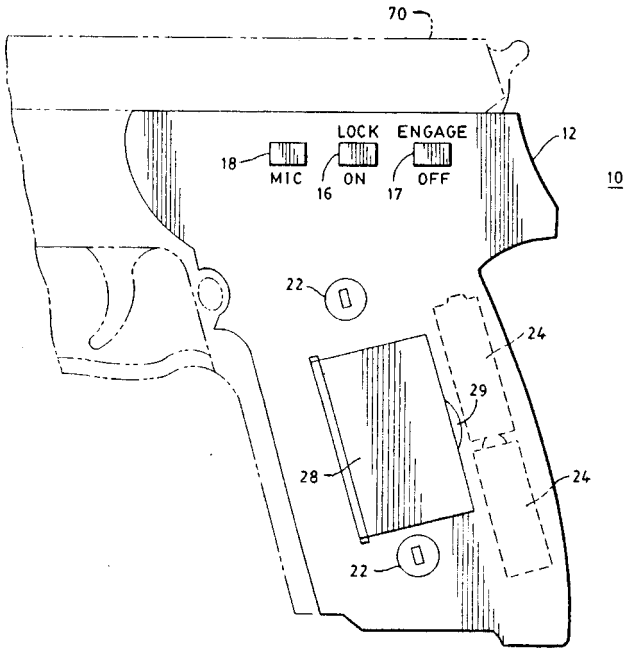
The patent presents a method for computing filler sounds to bridge the gap in case received speech samples are lost by a communications receiver due to channel errors or lost packets. First, a voiced/voiceless measure for the gap is estimated based on the recent history. Voiced filler is generated by extrapolating the latest pitch and linear prediction parameters. The prediction coefficients are modified such that formant bandwidths increase gradually during the filled interval. Voiceless gaps are filled by extending the noise output with a sound having similar correlation characteristics.—DLR

5,570,528

43.72.Ne VOICE ACTIVATED WEAPON LOCK APPARATUS

James W. Teetzel, Portsmouth, NH
5 November 1996 (Class 42/70.11); filed 14 July 1994

This apparatus could use any of several commercially available speaker recognition circuit chips for the task of verifying the user's identity before allowing that person to fire a weapon, such as a handgun or rifle which had been fitted with the device. The electronics unit is constructed on



a flexible plastic material, which may be stuffed into existing gun handles as well as fitted to new designs. A microphone in the handle grip material picks up the user's voice, passes it to the electronics unit which, if the match is successful, operates a solenoid which releases the firing mechanism.—DLR

5,572,570

43.72.Ne TELECOMMUNICATION SYSTEM TESTER WITH VOICE RECOGNITION CAPABILITY

John F. Kuenzig, assignor to Teradyne, Incorporated
5 November 1996 (Class 379/1); filed 11 October 1994

The widespread growth in the attachment of computer systems to the telephone network has been paralleled by an increase in the need for system test equipment beyond merely determining that telephone users can converse successfully. In 1992 the assignee of this patent introduced an automated tester, capable of executing a programmed script to generate DTMF tones, voice utterances, and various simulated line conditions. However, the unit had to be manually programmed and was difficult to use. This patent discloses an automatically programmable test unit which can learn test sequences during manual usage and by observing the telephone line conditions during such usage.—DLR

5,572,624

43.72.Ne SPEECH RECOGNITION SYSTEM ACCOMMODATING DIFFERENT SOURCES

Vladimir Sejnoha, assignor to Kurzweil Applied Intelligence, Incorporated
5 November 1996 (Class 395/265); filed 24 January 1994

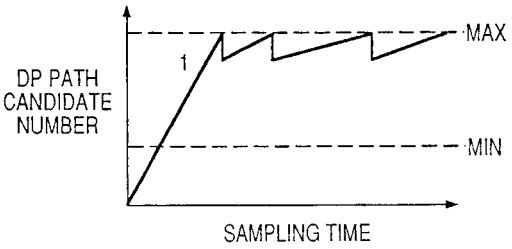
As the performance of large-vocabulary, continuous-speech recognition has improved in recent years, one factor that has not seen as much improvement is the ability to recognize male and female voices interchangeably. Most current systems require multiple speaker-description models, implying large demands on memory usage. This patent describes a transformation performed on the acoustic feature vectors from different genders, bringing them into a common space for the classification of phonetic units. Linear discriminant analysis is applied to the feature vectors, producing matrices that allow the minimization of within-class variance and the maximization of between-class separation for each phonetic unit in the training data.—DLR

5,577,162

43.72.Ne DYNAMIC PROGRAMMING MATCHING SYSTEM FOR SPEECH RECOGNITION

Yasushi Yamazaki, assignor to Fujitsu Limited
19 November 1996 (Class 395/2.41); filed in Japan 26 July 1991

As newer computing hardware provides ever-increasing computational capabilities, is it possible that we may see a resurgence of the almost forgotten methods of dynamic programming (DP) for speech recognition? This patent shows several optimizations which are applied to well-known DP techniques. The central feature is a system for monitoring the growth of



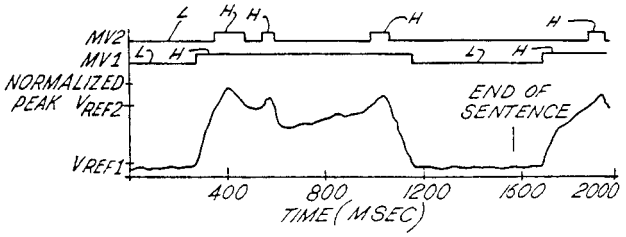
excessive path branching and of accordingly pruning the path tree. A branching threshold value is dynamically adjusted depending on the current number of paths, the goodness of fit along current paths, and the current percentage elapsed of the total utterance duration.—DLR

5,577,163

43.72.Ne SYSTEM FOR RECOGNIZING OR COUNTING SPOKEN ITEMIZED EXPRESSIONS

Peter F. Theis, McHenry, IL
19 November 1996 (Class 395/2.42); filed 21 September 1990

The patent describes a simple, low-cost approach, short of actual recognition, for classifying telephone utterances with the goal of improving the quality of information collected by an automated telephone query system. Based on a system of timers and measures of the speech energy envelope



pattern, the system attempts to locate stressed syllables and recurring patterns in the received utterance. The system is said to be able to distinguish items spoken as a list (including counting the items), and to detect sequences of spoken digits such as a phone number or zip code.—DLR

5,577,164

43.72.Ne INCORRECT VOICE COMMAND RECOGNITION PREVENTION AND RECOVERY PROCESSING METHOD AND APPARATUS

Kazue Kaneko and Keiichi Sakai, assignors to Canon Kabushiki Kaisha
19 November 1996 (Class 395/2.84); filed in Japan 28 January 1994

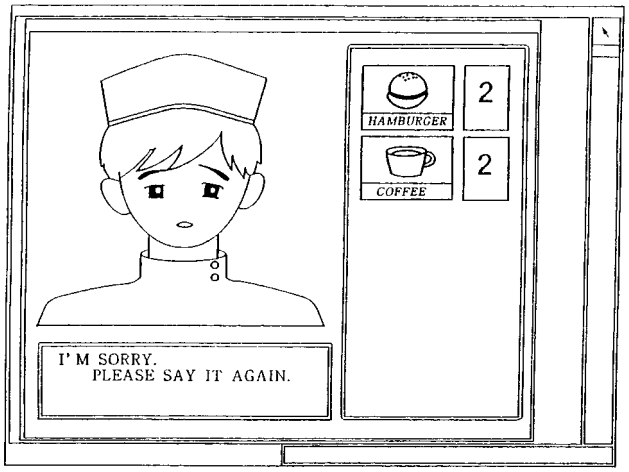
The patent describes a method for reducing false word recognitions in a small-vocabulary, discrete word recognizer. The method presented would seem to have the same effect as the widely used techniques of utterance-by-utterance dynamic syntax selection, while being much less general. In this case, a counter of allowable repetitions for each item is maintained in the dictionary.—DLR

5,577,165

43.72.Ne SPEECH DIALOGUE SYSTEM FOR FACILITATING IMPROVED HUMAN-COMPUTER INTERACTION

Yoichi Takebayashi *et al.*, assignors to Kabushiki Kaisha Toshiba
19 November 1996 (Class 395/2.84); filed in Japan 18 November 1991

This interactive speech dialogue system maintains an internal state representation, both for itself and for the human user, based on keywords and phrases recognized in the incoming speech. The application context is on-line ordering of items from a fast food menu. The communication also



includes a video link from which the computer can extract lip motions and by which a series of cartoon characters indicate current system states to the human user.—DLR

5,606,144

43.75.Bc METHOD OF AND APPARATUS FOR COMPUTER-AIDED GENERATION OF VARIATIONS OF A SEQUENCE OF SYMBOLS, SUCH AS A MUSICAL PIECE, OR OTHER DATA, CHARACTER OR IMAGE SEQUENCES

Diana Dabby, Cambridge, MA
25 February 1987 (Class 84/649); filed 6 June 1994

This patent describes a method for computer generation of different variations on a musical theme or piece, based on the properties of a chaotic system. An algorithm is shown to produce variations on a Bach Prelude as an illustrative example.—DWM

5,602,352

43.75.Gh VIBRATO ASSEMBLY AND ACOUSTIC COUPLING SYSTEM FOR STRINGED INSTRUMENTS

Richard E. Huff, Belmont, CA
11 February 1997 (Class 84/291); filed 24 July 1995

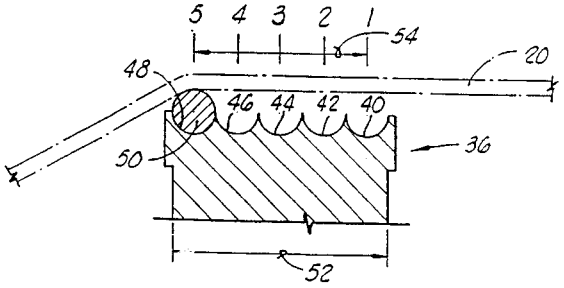
An acoustic coupling plate is described which can be divided into two plates, one of which couples the guitar body to the bridge, and the other couples the neck to the body. Flexure bearings for the assembly can be used to allow periodic variations of the tension of the strings for vibrato effects. These bearings are said to ''have the advantages of high strength, zero operational noise and rumble, and virtually zero friction and hysteresis.'' The ''Background of the Invention'' part of the patent document has a detailed description for the production of guitar tone and for previous vibrato mechanisms.—DWM

5,602,353

43.75.Gh BRIDGE SADDLE WITH ADJUSTABLE INTONATION SYSTEM

Henry E. Juszkiewicz, Nashville, TN and Timothy P. Shaw, Madison, TN
11 February 1997 (Class 84/298); filed 4 January 1995

The bridge saddle of an acoustic guitar is typically slanted, not exactly perpendicular to the center line of the instrument, in order to compensate for the effects of different string sizes, etc., upon the relative tuning of the strings. When string types are changed, the optimum termination points may



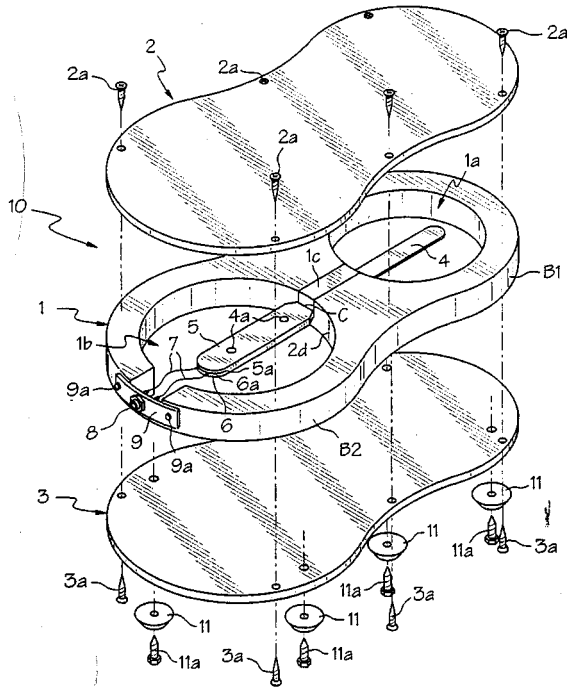
change. In this patent the bridge saddle has five parallel longitudinal grooves **40**–**48** so that the termination point can be changed to provide the best tuning for the set of strings employed.—DWM

5,602,354

43.75.Hi ACOUSTICAL RHYTHM BOARD

Thomas E. Martin, Cincinnati, OH
11 February 1997 (Class 84/410); filed 2 March 1995

The rhythm board described in this patent may be used by being struck by a musician's foot in accompaniment to the sound produced by the musician's musical instrument. The hollow rhythm board **1** is covered by wooden top **2** and bottom **3** which rests on rubber feet **11**. Freely resonant reed **4** in acoustic chamber **1a** extends into acoustic chamber **1b**, and the



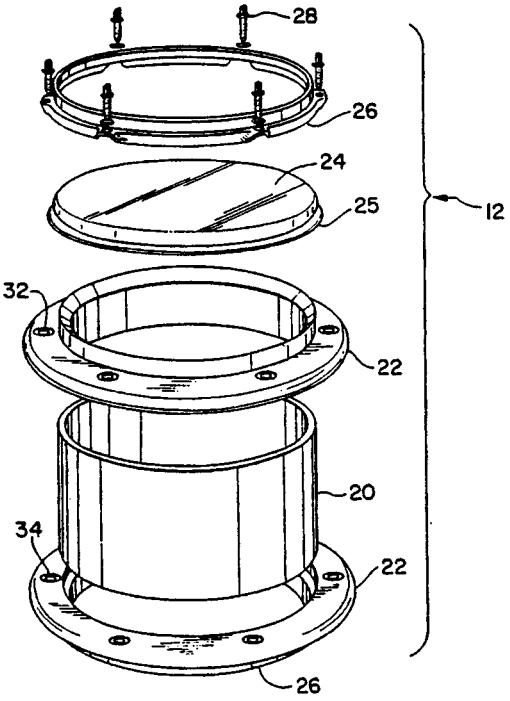
extension is enclosed in wood. A piezoelectric transducer **6** is also enclosed in wood, to receive the resonator vibrations. The signal output can be amplified for reproduction, or can be used to trigger electronically produced sounds in coordination with foot tapping by the musician.—DWM

5,606,142

43.75.Hi SHELL RESONANT MEMBRANOPHONE

Steven Volpp, assignor to Peavey Electronics Corporation
25 February 1997 (Class 84/411 R); filed 30 June 1994

The design of this drum frees the shell from the tension mechanism (attached only to a bridge) and from mounting hardware (also attached only to the bridge), allowing the shell to resonate more effectively. This is in



contrast to other drum patents reviewed recently in which the design purposely provided mechanical damping for the drum shell, to prevent undesired resonances.—DWM

5,612,502

43.75.Mn KEYBOARD MUSICAL INSTRUMENT ESTIMATING HAMMER IMPACT AND TIMING FOR TONE-GENERATION FROM ONE OF HAMMER MOTION AND KEY MOTION

Tomoyuki Ura, assignor to Yamaha Corporation
18 March 1997 (Class 84/687); filed in Japan 1 August 1994

In this combined acoustic and electronic piano, in which the player can choose between acoustic and electronic or combine them, the electronic system is equipped with key and hammer sensors so that their respective outputs can control the amplitude, timing, and frequency of the electronic signals in accordance with the musician's fingering on the keyboard.—DWM

5,602,356

43.75.Tv ELECTRONIC MUSICAL INSTRUMENT WITH SAMPLING AND COMPARISON OF PERFORMANCE DATA

Bernard Mohrbacher, assignor to Franklin N. Eventoff, Ferndale, WA
11 February 1997 (Class 84/609); filed 5 April 1994

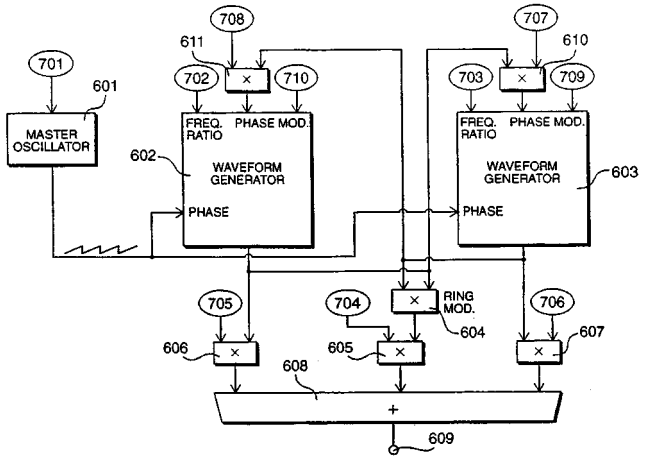
This ''electronic musical instrument'' is disclosed by a highly complex patent (26 pages of text and 28 figures) describing a system that enables a musical amateur or student to produce ''a relatively high quality individualistic rendition'' of an original live performance by a popular musician. The music system includes ''a keyboard strummer, in which the musical notes produced by the playing of the instrument are controlled by musical assistance data mapped unto the instrument keys and strum vanes from tracks specially prepared and synchronized with a prior performance of the piece. Modified mass media, such as CD ROM, TV signals and video cassettes are provided including synchronized note assist data and additional media, such as ROM packs or tone encoded audio cassettes or CDs, are provided with synchronizable note assist data for use with unmodified mass media.''—DWM

5,604,323

43.75.Tv MUSICAL TONE ELECTRONIC SYNTHESIZER AND METHOD

Anthony R. Hardie-Bick, assignor to Ethymonics Limited
18 February 1997 (Class 84/604); filed in the United Kingdom 7 September 1993

The Background section of this patent reviews the advantages of and the problems facing the digital methods of musical tone generation. Advantages include reduced cost, ease of manufacture, and the ability to reproduce the waveforms of real musical instrument tones. Disadvantages include the amount of memory required to store samples of a large number of waveforms for many fundamental frequencies, and to provide for tonal variation with dynamic changes. The mathematical modeling of the physical processes in musical tone generation, disclosed in recent patents, shows considerable promise but instruments based on these principles are said to be uneconomical until digital signal processing costs are reduced further. The



patent then discloses a synthesizing system having ''digital waveform generators **602**,**603** arranged such that a proportion of an output from a waveform generator **602** is supplied to a phase modulating input of a next waveform generator **603**, wherein a proportion of the output from the last waveform generator in a chain **603** of serially connected waveform generators is supplied to a phase modulating input of the first waveform generator **602** in said chain, thereby providing a degree of recursion. Outputs from a plurality of waveform generators are supplied to a ring modulator **604** for generating harmonic sums and differences of frequencies produced by said recursion.''—DWM

5,590,657

43.80.Sh PHASED ARRAY ULTRASOUND SYSTEM AND METHOD FOR CARDIAC ABLATION

Charles A. Cain *et al.*, assignors to The University of Michigan
7 January 1997 (Class 128/660.03); filed 6 November 1995

This system and method include a phased array that produces a focused ultrasonic beam of sufficient energy to ablate cardiac tissue. In focusing the beam, aberrations produced by transmission through inhomogeneous tissue between the array and the treatment volume are taken into account by sensing the phase distribution caused by the aberrations and calculating a compensating phase distribution that is employed in the excitation of the array. The system also permits real time correction of the beam position to follow a moving volume of myocardial tissue.—RCW

5,590,658

43.80.Vj PORTABLE ULTRASOUND IMAGING SYSTEM

Alice M. Chiang and Steven R. Broadstone, assignors to TeraTech Corporation
7 January 1997 (Class 128/661.01); filed 29 June 1995

This system consists of a handheld scan head coupled by a cable to a battery-powered data processor and display unit that is preferably a lap-top computer. The scan head enclosure houses, in addition to an array of ultrasonic transducers, circuitry that includes a pulse synchronizer used in transmission and a beamformer used in receive. The electronics employ low-power integrated circuits, and signal loss is reduced by the proximity of the electronics to the array. The cable used to transfer data between the scan head and the data processor and display unit is less complicated than in systems that transfer all of the transducer signals to a console signal processing unit.—RCW

5,590,659

43.80.Vj ULTRASONIC VISUALIZATION METHOD AND APPARATUS

Robin Hamilton *et al.*, assignors to Intravascular Research Limited
7 January 1997 (Class 128/661.01); filed 17 July 1995

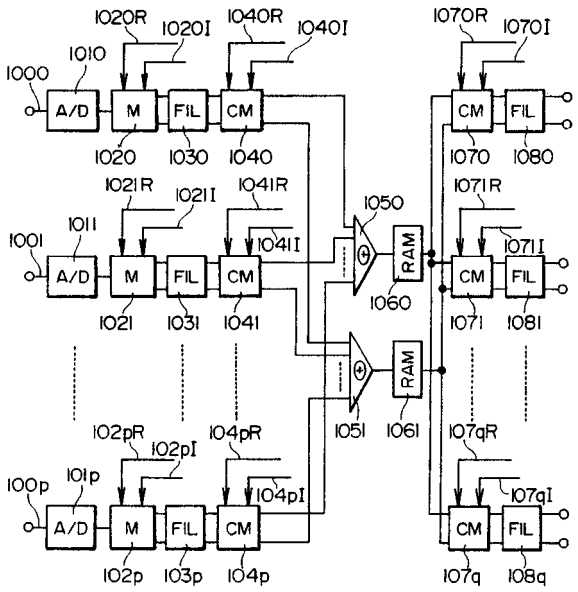
In this method and apparatus, each element or group of elements in a transducer array is connected using a controller that determines the order in which the elements are energized for transmission and the order in which they receive echos. Beam profiles with low sidelobes are produced by a synthetic aperture technique in which different combinations of elements are used for transmission and reception. Beam profiles are also improved by an optimal filtering technique using a more complex approach than employed before in this application.—RCW

5,591,911

43.80.Vj ULTRASOUND SIGNAL PROCESSOR

Hiroshi Masuzawa *et al.*, assignors to Hitachi Medical Corporation
7 January 1997 (Class 73/602); filed in Japan 5 July 1995

This signal processor has a parallel structure in which baseband signal processing is accomplished digitally using A/D converters **101p**, complex multipliers **102p**, low-pass filters **103p**, complex multipliers **104p**, adders



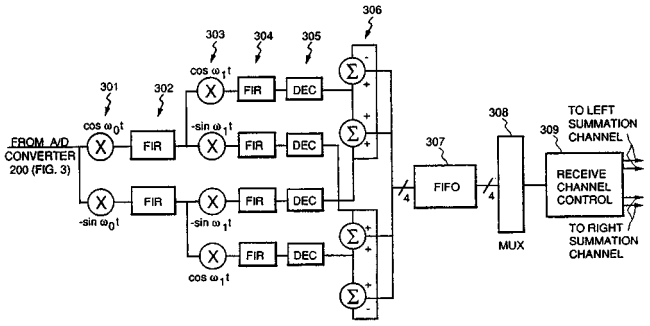
1050 and **1051**, memories **1060** and **1061**, complex multipliers **107q**, and low-pass filters **108q**, arranged as shown above.—RCW

5,600,675

43.80.Vj ULTRASONIC IMAGER HAVING IMRPOVED BANDWIDTH

William E. Engler, assignor to General Electric Company
4 February 1997 (Class 375/261); filed 7 September 1994

Improved bandwidth is achieved by employing two demodulation stages in each receive channel. Inphase and quadrature components from the first demodulator stage are also demodulated in the second stage to produce a total of four signals that are filtered and combined to form two sets of complex signals. One set contains the upper frequencies in the receive signal



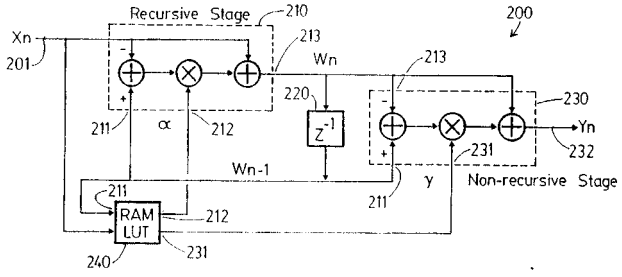
while the other contains the lower frequencies. Each band is separately phase rotated using the center frequency of the corresponding subband. Upper-band signals and lower-band signals are next coherently summed, and then the upper-band and lower-band signals are reconstituted into a single baseband signal.—RCW

5,595,179

43.80.Vj ADAPTIVE PERSISTENCE PROCESSING

J. Nelson Wright *et al.*, assignors to Acuson Corporation
21 January 1997 (Class 128/660.07); filed 2 May 1995

In this processing, a recursive filtering stage is followed by a nonrecursive filtering stage. Coefficients for each stage come from a look-up table that is addressed using input or output pixel values. A recursive filtering coefficient α is held constant over time while a nonrecursive filtering coefficient γ is adjusted dynamically. The nonrecursive filter stage has a frequency response that attenuates high-frequency components. The response time of the nonrecursive stage only spans one frame interval while the response time of the recursive stage is generally longer and dependent on the recursive filtering coefficient.—RCW



Biomimetic sonar recognizes objects using binaural information[a)]

Roman Kuc
Intelligent Sensors Laboratory, Department of Electrical Engineering, Yale University, New Haven, Connecticut 06520-8284

(Received 17 June 1996; accepted for publication 12 February 1997)

An active sonar using Polaroid electrostatic transducers positioned at the end of a robot arm is described that adaptively changes its location and configuration in response to the echoes it observes in order to recognize an object. The sonar mimics biological echo-location systems, such as those employed by bats and dolphins, in that there is a center transmitter flanked by two adjustable receivers, the sonar has rotational and translational mobility, and the echo processing contains elements that have been observed in the mammalian auditory system. Using information in the echoes, the sonar translates in a horizontal plane and rotates in pitch and yaw to position an object at a standard location within the beam patterns. The transmitter points at the object to maximize the incident acoustic intensity and the receivers rotate to maximize the echo amplitude and the bandwidth, and to minimize the echo-producing region. This procedure results in a unique echo vector associated with each object at a given object pose. Recognition is accomplished by extracting 32 values from the binaural echo patterns and searching a data base that is constructed during a learning phase. The system operation is illustrated by having it recognize rubber O-rings of different sizes and by differentiating the head and tail sides of a coin. © *1997 Acoustical Society of America.* [S0001-4966(97)05506-9]

PACS numbers: 43.10.Ln, 43.80.Jz, 43.80.Nd, 43.80.Ka [FD]

INTRODUCTION

Bats and dolphins employ echo location for prey identification by emitting a series of acoustic pulses and processing the echoes.[1–3] This paper combines the techniques employed by biological systems with principles from acoustics, neuroscience, robotics, and signal processing to implement a sonar system that can recognize objects from their echo waveforms. The sonar is adaptive in that it changes its configuration using the information in the detected echoes. The sonar design is motivated by bats, whose ears react by rotating to the direction of the echo source, and by dolphins, who appear to move as if to position the object at a standard location in three-dimensional space.

Manmade sonar has had limited success for object recognition because the echo waveform varies significantly with the location of the object within the acoustic beam pattern. This variation was minimized previously by careful placement of the objects within the acoustic field and by using objects that are large compared to the wavelength and that have shapes (spheres, cubes, pyramids, and cones) that exhibit very different scattering properties.[4–7] The system described in this paper exploits the important biological principle of sensor mobility by translating and rotating the sonar in order to standardize the view of the object by positioning it at a constant location. The complexity of the detected echo is then reduced, thus making object recognition from the echo waveforms feasible.

[a)]''Selected research articles'' are ones chosen occasionally by the Editor-in-Chief, that are judged (a) to have a subject of wide acoustical interest, and (b) to be written for understanding by broad acoustical readership.

I. BIOMIMETIC SONAR SYSTEM

Modeled after biological systems, our sonar shown in Fig. 1 consists of three Polaroid electrostatic transducers that can act as either transmitters or receivers.[8] This transducer is similar to those used in the Polaroid autofocus camera and acoustic digital tape measures. The transducer has a 3.75-cm-diam circular aperture and exhibits a broad resonance at 60 kHz with a nominal wavelength in air (c=343 m/s) equal to λ=5.7 mm. The beam has a full-width half-maximum amplitude spread of 10°. The protective cover was removed, exposing the foil surface of the transducer, to eliminate internal reflections that limit the bandwidth.

A center transmitter (mouth) is flanked by two receivers (ears), with the separation between the centers of the mouth and each ear being equal to D=4.5 cm. The ears are rotated in opposite directions by cables connected to a stepper motor to accomplish a type of focusing operation. The entire assembly is mounted on the end of a small educational robot arm whose end has three translational (x, y, z) and two rotational (sonar pitch and yaw) degrees of freedom. The sonar can then be translated and rotated so that it can view an object from any desired aspect.

The sonar moves in a horizontal plane 11 cm above a smooth working surface and is directed downward at a pitch of −45°. Using such a pitch directs the beam reflected from the surface away from the ears, leaving only the information-bearing echoes scattered by the object to be detected by the ears.

Initially, the mouth and both ears point in the same direction to maximize the echo-producing region for object detection over a large field of view. The environment is scanned by transmitting interrogation pulses and processing the signals detected at the ears while rotating the sonar about

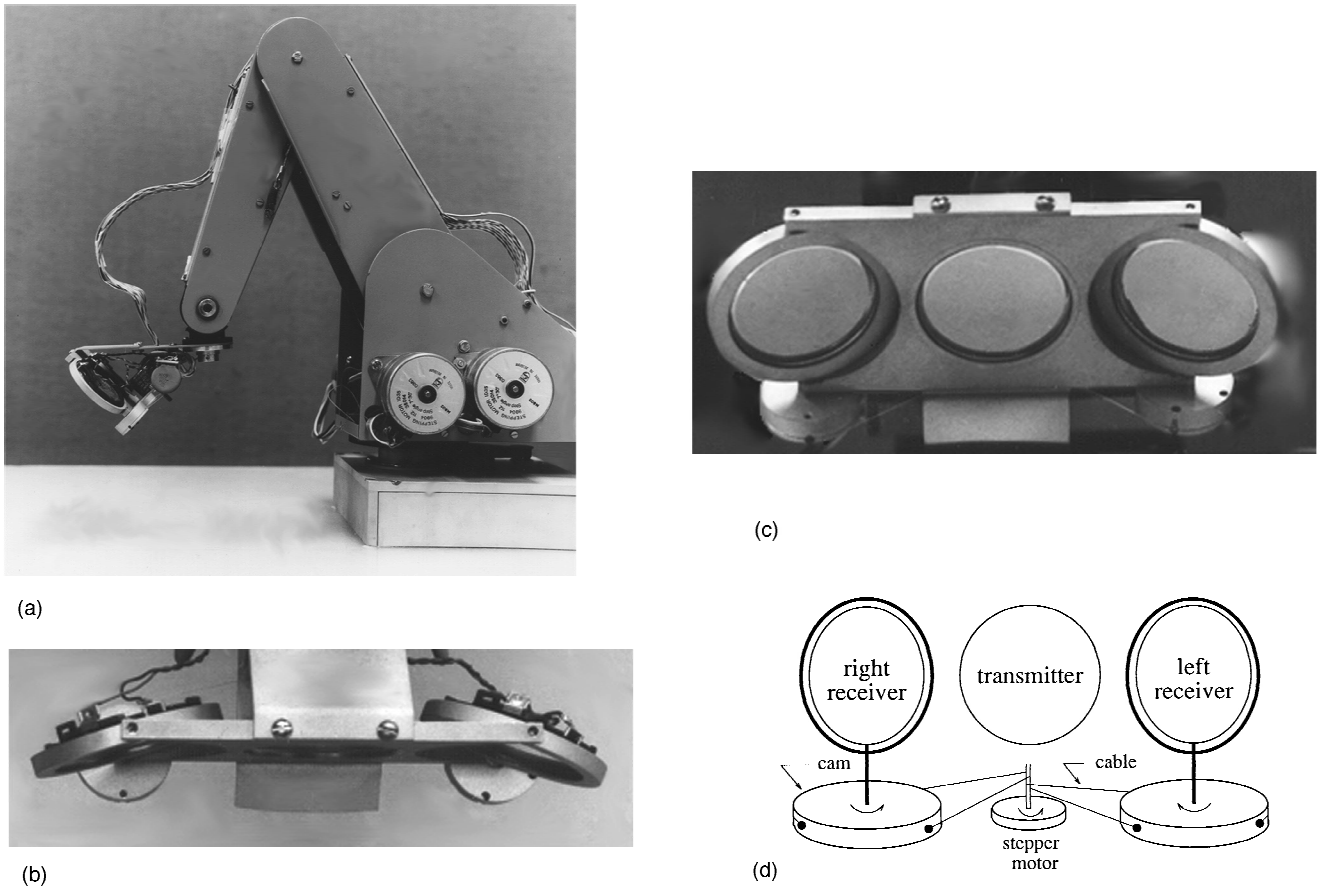


FIG. 1. Biomimetic sonar system. (a) Sonar situated at the end of the Armdroid 1 robotic arm, manufactured by Colne Robotics Co. Ltd. in England. The arm consists of two segments each 19 cm long. (b) Top view. (c) Front view. (d) Block diagram. The sonar consists of three Polaroid electrostatic transducers configured as a center transmitter flanked by two receivers that can rotate to focus on an echo-producing object.

the yaw (vertical) axis to sweep across the environment. The signal at each ear is acquired by using one of the two channels of an analog-to-digital converter (Gage CS1012) sampling at 2 MHz per channel with 12-bit resolution. A new interrogation pulse is transmitted as soon as the echoes have been detected and processed. Being controlled by a Pentium 120 processor, pulses are emitted as often as every 100 ms. Once echoes are detected, the sonar moves to position the object along the transmitter axis at a range of 15 cm. The ears are then rotated to place the reflecting part of the object on their axes. The 15-cm range is in the far field of each transducer ($r > a^2/\lambda$), where a is the radius of the transducer), thus avoiding the spatial variation of the near-field wavefronts, but is in the near field of the three transducer system ($r < D^2/\lambda$), thus obtaining two novel views of the object.

II. ECHO PROCESSING

An interrogation pulse is produced by impulsively exciting the transmitter. The echoes detected at each ear are processed to extract the time-of-flight (TOF) for adjusting the yaw and focus and the envelope information for object recognition. The waveform and spectrum of an echo from a plane at 0.5-m range is shown in Fig. 2. The waveform has a short duration or, equivalently, a wide frequency bandwidth, making it useful for differentiating nearly identical objects.

The TOF is determined from the time that the waveform samples first exceed a threshold. The TOF from each ear defines an ellipse of possible object locations, with the ear and mouth being the foci. The object location in the plane containing the transducer axes, in terms of range r and azimuth θ, is determined from the intersection of the two ellipses in the forward direction.[9] Using sonar feedback, the system maintains the range at 15 cm to within ± 0.1 mm. The object is positioned on the transmitter axis by driving θ to zero, as shown in Fig. 3. The accuracy of $\pm 0.1°$, limited by the resolution of the robot stepper motors. While the object is being driven toward the transmitter axis, the receivers are rotated so that their axes point toward the object using the angle

$$\phi = \arctan\,(r/D).$$

After the object has been positioned along the axis of the mouth at a range of 15 cm and the ears are focused, the system can either learn a new object by acquiring echoes and including them in a data base (memory) or recognize an object that is already included in the memory. In the learning phase, the sonar pitch is varied to observe the object at a variety of elevations. Echoes are obtained during a scan in elevation as the pitch is varied in 0.225° steps about the nominal pitch value from $-43°$ to $-47°$.

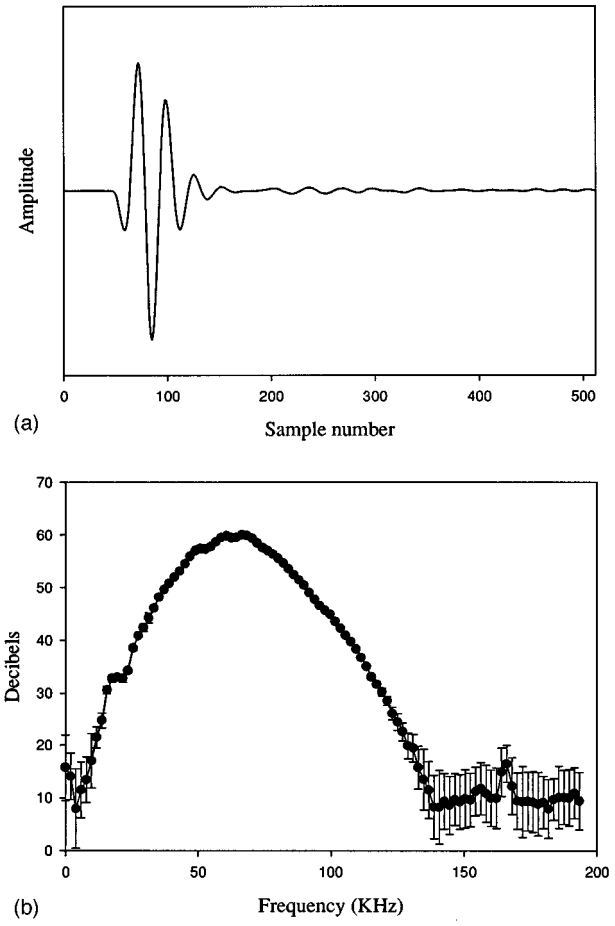


FIG. 2. Echo from plane reflector located at a range of 50 cm. (a) Echo waveform. (b) Power spectral density average and one standard deviation interval.

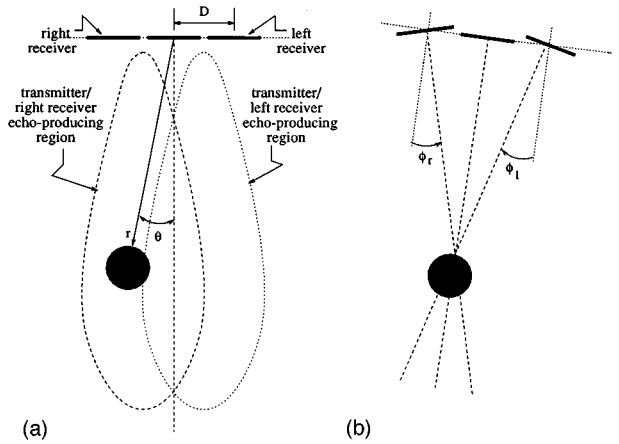


FIG. 3. View of plane of sonar beams in unfocused and focused configurations. An object initially lies at range r and angle θ relative to the transmitter. (a) Unfocused configuration. The echo-producing regions indicate the locus of points that produce detectable echoes. The object is located such that only the right receiver detects an echo. (b) Focused configuration. The sonar rotates to position the object on the transmitter axis and the angles ϕ_r and ϕ_l position the object on the right and left receiver axes, respectively.

The steps in echo processing that occur in each ear are shown in Fig. 4. The echo waveform contains 512 samples, starting 50 samples prior to the time the threshold is exceeded to allow low level signals at the start of the echo packet to be included in the analysis. This data window duration (512) was chosen to include all the echoes from the largest object in our collection. The waveform data are rectified and logarithmically compressed. The logarithmic compression reduces the dynamic range of the echoes and emphasizes the important small-amplitude structure that is present in the echo packet. The log-rectified signal is then applied to a second-order low-pass filter whose frequency response falls to -3 dB at 12 kHz to produce the echo envelope. Since the envelope exhibits a slow variation in time, its 512 samples can be reduced to a smaller number without losing significant information. This is done by segmenting the envelope into sixteen 32-sample segments and computing the mean for each segment. This segment size corresponds to approximately one period at the resonant frequency of the transducer. This data reduction allows the echo packet to be represented by an *envelope vector* containing 16 values. The envelope vector is then normalized to have a constant energy. Finally, an object is characterized by combining the vectors from each ear to produce a 32-value vector. In the learning phase an object is scanned and the envelope vectors are stored in memory. The memory contains the object identity, its pose, and the envelope vectors from the left and right ears.

During the recognition phase, an object is viewed at a nominal pitch of $-45°$, but with an error of $\pm 1.5°$ due to errors in the joints. The sonar examines the object using a single interrogation pulse. The observed echoes are processed to extract the envelope vectors and these are then compared to those stored in memory. The accuracy of the robot arm was sufficient to assure that the pitch of the sensor was within the interval used in the learning stage and that a corresponding envelope vector, to within a noise tolerance, resided within the memory. The observed envelope vector pair was compared against those in memory by computing the squared error. The vector pair in memory that produced the minimum error was used to indicate the object identity.

The echoes typically contain random fluctuations whose

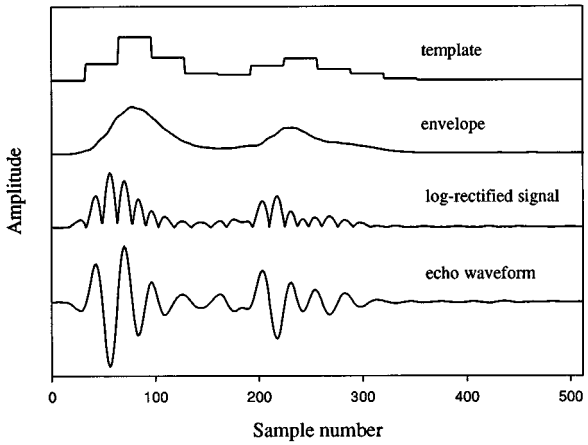


FIG. 4. Processing of echoes. From bottom to top, the waveforms indicate an echo packet from an object, the rectified waveform after logarithmic compression, the envelope, and a plot of the 16 vector values that form half of the template used to represent objects.

effects on the vector pair must be reduced to differentiate similar objects. The sources of these fluctuations include random amplitude fluctuations due to temperature and density variations in the medium between the transducers and object, thermal noise, and sampling jitter in acquiring the echo waveform. These fluctuations not only produce errors in the echo envelope magnitude, but also in its time registration, since it is determined from the time the echo waveform first exceeds the threshold. In the learning stage, 50 echoes are processed at each increment in pitch angle. Each observed envelope vector pair is considered to be a candidate and compared to those vector pairs already residing in memory associated with the particular object. If the squared error between the candidate and any vector in memory is less than a *perceptual threshold* then the candidate is discarded, since a representative vector already exists in memory. The perceptual threshold level is four times greater than the average error value produced by noise alone, and accounts for the variations in the propagation medium. On average, for every 50 candidates considered, approximately four are added to the memory. Hence, after the learning phase, each object is represented by approximately 50 vectors that form clusters, each cluster corresponding to a particular viewing pitch angle.

III. ECHO PRODUCTION

The mechanisms that cause an object to produce echoes are described by the model proposed by Freedman.[10,11] This model can be simply stated: Echoes are produced whenever an incident pulse encounters either a discontinuity in the acoustic impedance or a discontinuity in any of the derivatives of the acoustic impedance with respect to range. The echo magnitude is proportional to the size of the discontinuity and inversely related to the order of the derivative that exhibits the discontinuity. The magnitude is also related to the location of the discontinuity in the transmitting and receiving beam patterns. A typical object will contain echo-producing features that are distributed over range, azimuth, and elevation. Only the feature at closest range that produces a detectable echo is employed to position the sonar. Even if different features are detected by each ear, the object is positioned near the transmitter axis at a repeatable location.[9]

For example, consider the echoes observed from a rubber O-ring. An O-ring lying on a surface producing echoes that are similar to those from a hoop floating in the water as detected by a dolphin. The echo-producing features are shown in Fig. 5. The echo waveforms observed from three O-rings of different sizes are shown in Fig. 6. Since O-rings are symmetric with respect to angular rotations, identical waveforms are observed at both ears. One of the envelope vectors for each O-ring is shown in Fig. 7. The system can easily differentiate the three different O-rings without error.[12]

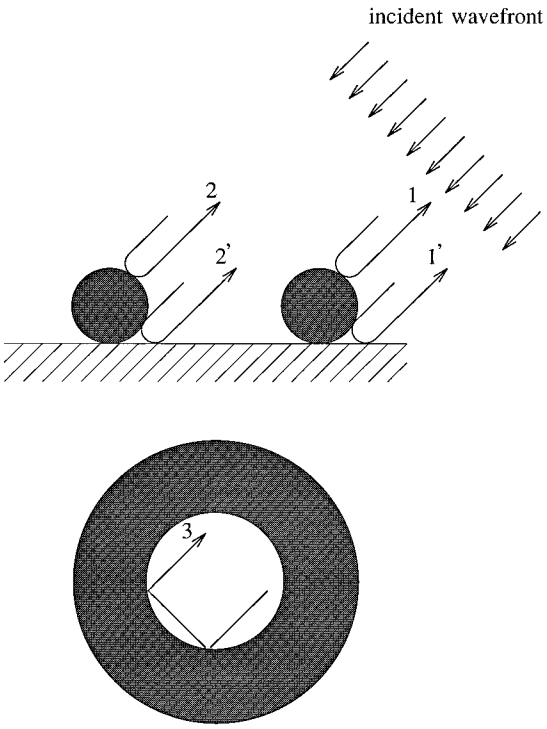


FIG. 5. Paths of echoes reflected from O-rings. The first echo detected in time is indicated by a 1, with the echo 1′ following closely after, followed by the pair 2-2′. The echo path labeled 3 is a multiple reflection from the inner surfaces.

IV. DIFFERENTIATING HEADS FROM TAILS OF A COIN

To illustrate the sensitivity of our sonar we attempt to differentiate the head from tail side of a coin. The echo packet observed from a coin consists primarily of two components, as shown in Fig. 8.

(1) The first component, corresponding to the discontinuity in the acoustic impedance, is a reflection from a patch of the coin edge to the table surface and back to the receiver. A second path also exists from the table surface to the coin edge and back to the receiver, doubling the magnitude of the echo. Smaller first-order echoes occur from the corners in the relief patterns, such as those in the leading edges of the lettering.
(2) The second component corresponds to the discontinuity in the first derivative of the acoustic impedance that can take on either positive or negative values. Echoes from negative discontinuities experience a 180° phase reversal. These are indicated in dashed lines in the figure. These second-order echoes occur primarily at the sharp edges around the top surface of the coin, but also from features in the relief pattern. Similar echoes have been observed from a disk reflector by Simmons.[13]

The coin used in the experiments was a 1996 U.S. penny in mint condition, a disk having 18.94-mm diameter and 1.50-mm thickness. In the learning state, the penny was positioned with Lincoln head upright and looking to the right. The system learned the coin by acquiring a set of approximately 50 vector pairs during the sweep in pitch. Because the

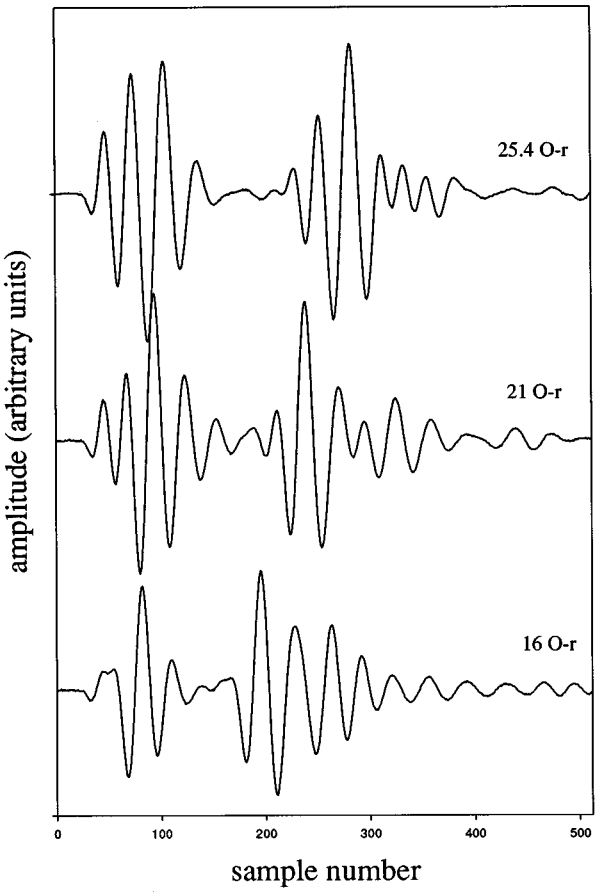


FIG. 6. Waveforms of echoes reflected from rubber O-rings observed at one receiver. O-ring data are denoted by ''O-r'' with the preceding number indicating the outside diameter in mm.

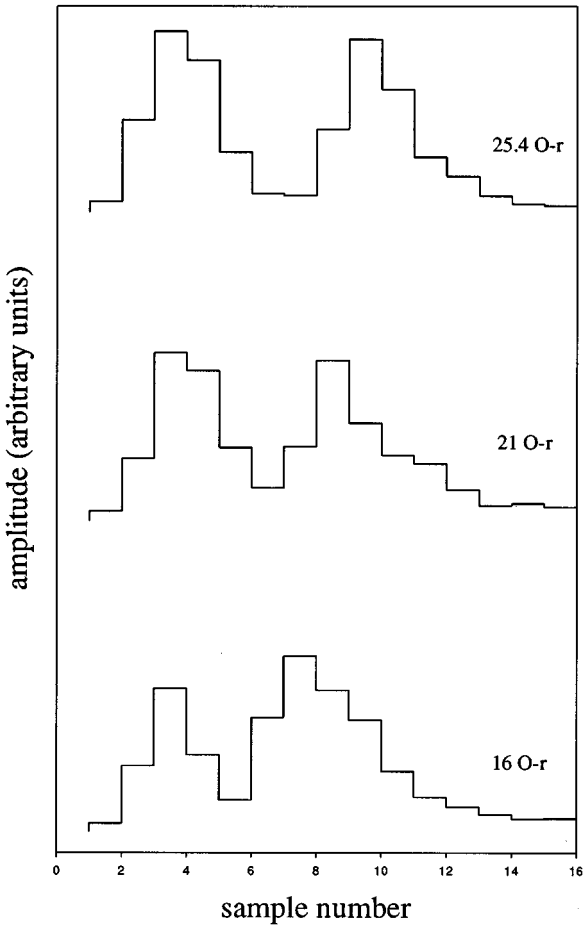


FIG. 7. Plots of the 16 values extracted from the echoes reflected from O-rings detected at one of the receivers. The data base consists of 32-element vectors containing the 16 values from the right and left receivers.

coin is not symmetric, different vector values are observed at each ear. Typical echo waveforms from the penny heads are shown in Fig. 9 and the corresponding envelope vectors are shown in Fig. 10. The vectors were normalized to a value that provides sufficient accuracy when represented by a three-digit integer in the memory.

Without moving the coin or the sonar, a series of recognition trials was conducted immediately after completion of the recognition phase. At each trial, the observed vectors were compared with those in memory. For this set of trials, the squared error was due to noise and sampling-jitter effects. The squared error values observed in 100 trails are shown in Fig. 11. The mean is 388 with a standard deviation of 86. This mean value represents an error of about 1% in the largest magnitude elements in the vector.

The penny was then flipped to expose the tail side, with the Lincoln Memorial upright, and placed at the same location within ± 0.1 mm. A set of 100 recognition trials was then repeated. Typical echo waveforms from the penny tails are shown in Figs. 9 and 10. At each trial, the observed penny-tail vector pair was compared with the penny-head vector pairs in memory. Because the coin size, of course, is the same, the squared error was due to the differences between the relief patterns and noise. The errors observed in 100 trials are shown in Fig. 11. The mean of the squared errors is 7644 with a standard deviation of 344. This result indicates that the echoes from the head and tail sides are significantly different.

The penny was then replaced with a 1996 U.S. dime showing a head, with Roosevelt looking to the left, in mint condition. The dime is a disk having 17.80-mm diameter and 1.35-mm thickness. Because the echo acquisition starts at the

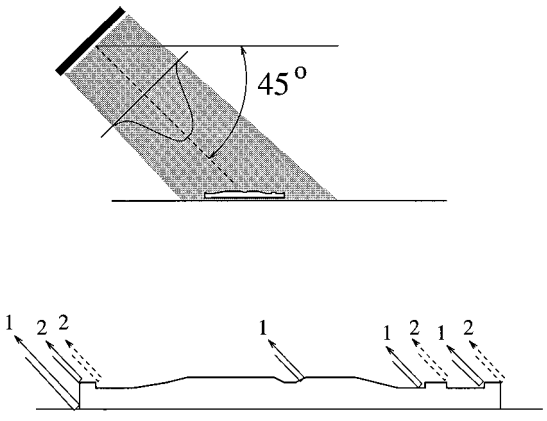


FIG. 8. Coin located in transmitter field. Two orders of echoes are observed: 1—echoes from discontinuities in acoustic impedance experienced by the propagating pulse, 2—echoes from discontinuities in the first derivative of the acoustic impedance with respect to range. Dashed lines indicate 180° phase shift in the echo.

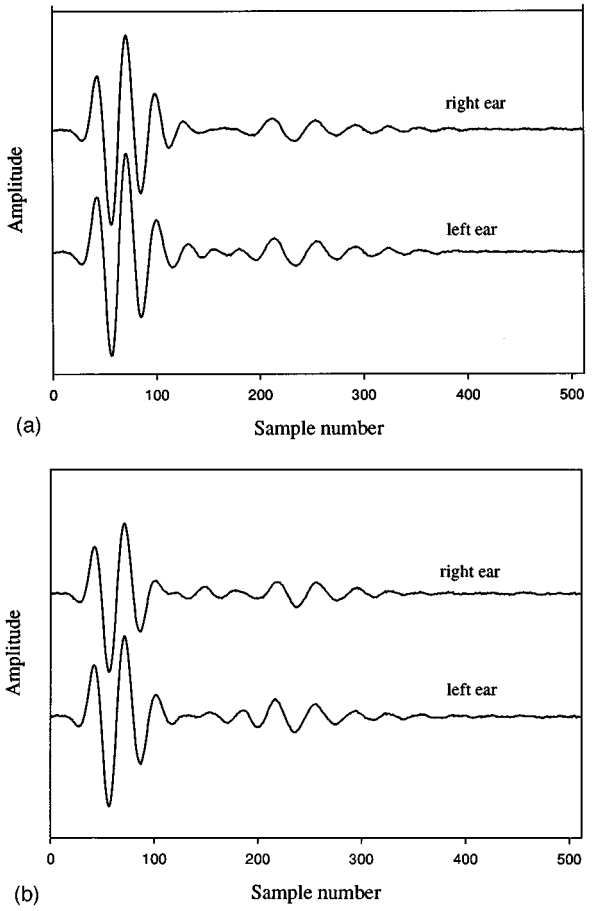


FIG. 9. Echo waveforms from penny coin. (a) Echoes from head side. (b) Echoes from tail side.

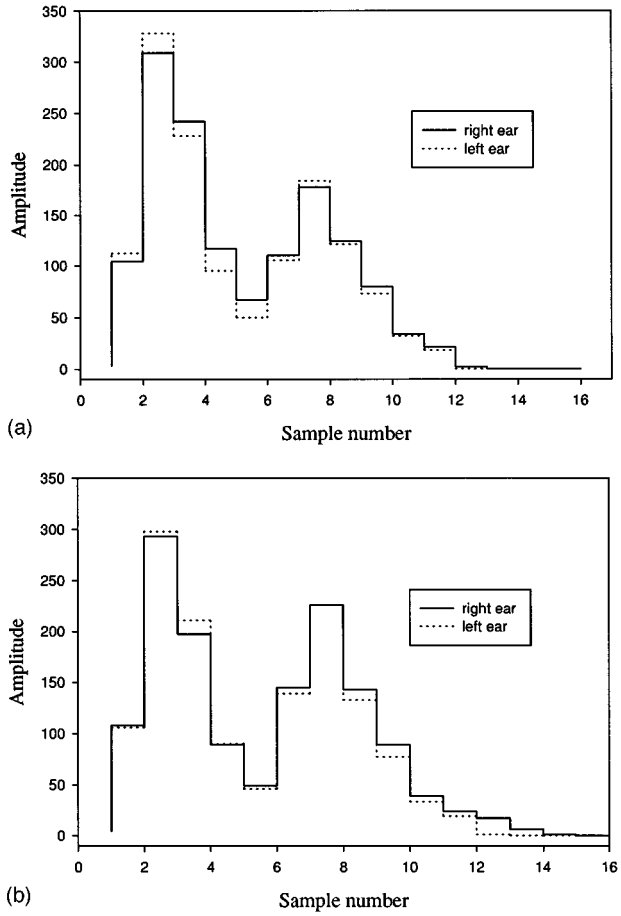


FIG. 10. Echo envelope vectors from penny coin. (a) Values from head side. (b) Values from tail side.

near edge of the coin and the dime is smaller, the distant edge echo from the dime occurs at a shorter delay than that of the penny. This results in large errors between the observed dime vectors and the penny vectors in memory. To illustrate this, a set of 100 recognition trials was conducted. At each trial, the observed dime-head vector pair was compared with the penny-head vectors in memory. The errors observed in 100 trials are shown in Fig. 11. The mean of the squared errors is 12 573 with a standard deviation of 391. This result indicates that the echoes from pennies and dimes are significantly different.

Since the relief pattern on a coin is not symmetric, the vector values are related to the pose of the coin. The pose was measured as a counterclockwise angular rotation from the upright Lincoln head. The pose of the penny-head was changed in 5° increments, and a set of 20 recognition trials was conducted at each pose value. At each trial, the observed penny-head pose vector pair was compared with the penny-head 0° templates in memory. The errors are shown in Fig. 12. The errors are observed to be pose dependent with magnitudes that can vary significantly. To recognize a penny-heads having an arbitrary pose, vectors from the different poses must be acquired in the learning phase. The pose must be varied in sufficiently small steps, such that the change in the vectors is comparable to the variation caused by the noise.

V. DISCUSSION

Object recognition occurs when the envelope vector pair observed during the recognition stage matches a vector pair in memory to within an acceptable error. If the minimum observed squared error is larger than that expected solely on the basis of noise, then it is assumed that a novel object has been observed. If the identity of this new object is made

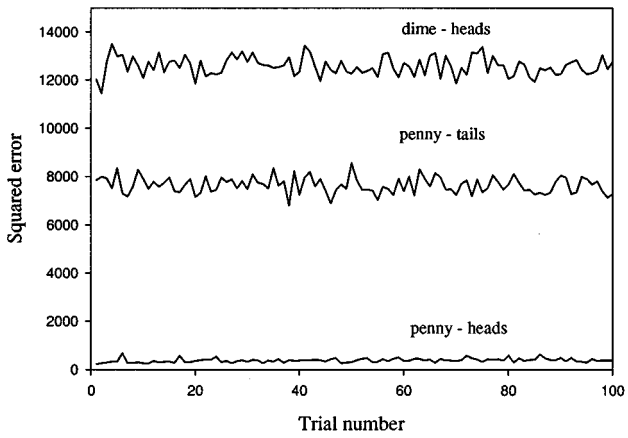


FIG. 11. Squared error between observed echo vectors and penny-head templates in memory for penny-heads, penny-tails, and dime-heads.

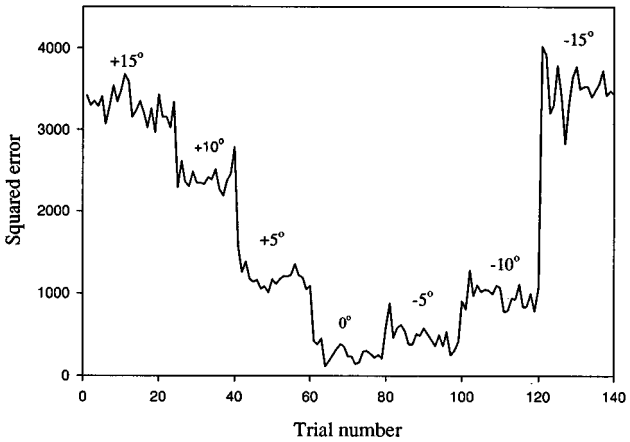


FIG. 12. Squared error as pose of penny-head is varied.

known to the system, the system reverts to the learning phase and increases the memory contents by including the vectors from the new object.

The echo data segment processed corresponds to a 512 sample data window. This duration is related to the size of the objects that were to be recognized. The window started just prior to the time the echo from the nearest feature exceeded a threshold to include any leading low level echoes, a sort of *pretriggering* operation. The 512 samples acquired at a 2-MHz sample rate represent a 0.25-ms window. At the sonic speed of 343 m/s, this time window corresponds to a 4.3-cm range interval. This allowed all the echoes produced by the 2.54-cm O-ring to be acquired. It would be interesting to determine if biological systems use an adaptive window duration related to the size of the intended object of search. That is, are large objects recognized by enlarging the data window duration, or by piecing together evidence derived from a collection of shorter windows?

The binaural nature of the sonar was used primarily for accurate positioning of the object in azimuth by adjusting the yaw of the sonar. The object recognition aspect exploits binaural hearing by obtaining two novel views of an object from one position of the sonar.

The penny-pose experiment indicates that to recognize a head in a pose-independent case, additional templates must be acquired at increments to pose less than 5°. The increment for a particular object depends on the complexity of its shape and its size relative to the beam cross section. Symmetric objects yield their identity from a single pose.

The cochlear model used in our echo processing is a modification of the classical bandpass filter bank, motivated by our sonar characteristics and the capabilities of our data acquisition system. We thought it would be an interesting exercise to consider such an evolutionarily parsimonious single hair-cell model whose frequency characteristics correspond to those of the transmitting/receiving transducers shown in Fig. 2.

The paper considers the problem of recognizing isolated objects that are small enough to be contained within the beam patterns. When more than one object is present, the echoes can interfere and the match between the observed vectors and those in the memory is poor. For example, let us consider two coins lying in close proximity. If the edge of one coin is resting on the other, the tilted coin would exhibit echo waveforms that are not in the memory that was formed using coins lying flat. In this case, the pair of coins may need to be considered as one object and learned in that manner. If a tolerance in the vector match is allowed, the set of such configurations is finite. Then, one possible solution is to treat each configuration as a separate object and learn its echo vectors using the procedures described above. Clearly, this approach is only the first step toward solving this problem.

A fundamental problem arises when scanning a large object, or one that is not entirely contained within the beam patterns. As the object pose is changed, a new echo-producing feature can enter the sonar beam and can cause the TOF to change significantly. Hence, the values of the envelope vector can change in a discontinuous manner with pose. This problem is an inherent weakness in relying exclusively on the TOF. Perhaps results from animal studies will provide some insights on how to deal with this problem.

VI. SUMMARY

This paper described an adaptive sonar system that is mounted on the end of a robot arm that recognizes objects using echo location. The sonar standardizes the echo waveform by being mobile and adaptively adjusting its configuration. The processing is biologically motivated in its mobility and configurational change. A memory is established and enlarged in the learning phase. A scan in pitch is performed to compensate for possible errors in the pitch angle during the recognition phase. The echoes are processed to produce a pair of envelope vectors. Multiple pulses are emitted and only the novel vectors, or those that exceed a perceptual threshold, are stored in memory. The recognition task is accomplished by having the system repeat the positioning operations and emit a single interrogation pulse. The echo vector values are extracted and the memory is searched to find the best match in the least squared error sense. Differentiating heads from tails of a coin illustrates the sensitivity of the sonar.

ACKNOWLEDGMENT

The research described in this paper was supported by the National Science Foundation under Grant No. IRI-9504079.

[1] P. E. Nachtigall and P. W. B. Moore, *Animal Sonar: Processes and Performance* (Plenum, New York, 1988).
[2] J. A. Simmons, M. B. Fenton, and M. J. O'Farrell, ''Echolocation and pursuit of prey by bats,'' Science **203**, 16–21 (1979).
[3] W. W. L. Au, *The Sonar of Dolphins* (Springer-Verlag, New York, 1993).
[4] J. M. Richardson, K. A. Marsh, D. Gjellum, and M. Lasher, ''Acoustic recognition of objects in robotics: Determination of type, pose, position, and orientation,'' in *Acoustic Imaging*, edited by L. Kessler (Plenum, New York, 1988), pp. 613–620.
[5] H. L. Roitblat, P. W. B. Moore, P. E. Nachtigall, R. H. Penner, and W. W. L. Au, ''Natural echolocation with an artificial network,'' Int. J. Neural Networks **1**, 239–248 (1989).
[6] I. E. Dror, M. Zagaeski, and C. F. Moss, ''Three-dimensional target recognition via sonar: A neural network model,'' Neural Networks **8**, 149–160 (1995).
[7] K. Sasaki and M. Takano, ''Classification of objects' surface by acoustic transfer function,'' in *Proceedings IEEE/RSJ International Conference on Intelligent Robots and Systems* (IEEE, Piscataway, NJ, 1992), pp. 821–828.

[8] C. Biber, S. Ellin, E. Sheck, and J. Stempeck, ''The Polaroid ultrasonic ranging system,'' in *Proceedings of the 67th Audio Engineering Society Convention,* 1990. Reprinted in Polaroid Ultrasonic Ranging System Handbook.
[9] R. Kuc, ''Biologically-motivated adaptive sonar system,'' J. Acoust. Soc. Am. **100**, 1849–1854 (1996).
[10] A. Freedman, ''A mechanism of acoustic echo formation,'' Acustica **12**, 10–21 (1962).
[11] A. Freedman, ''The high frequency echo structure of some simple body shapes,'' Acustica **12**, 61–70 (1962).
[12] R. Kuc, ''Fusing binaural sonar information for object recognition,'' in *Proceedings of the 1996 IEEE/SICE/RSJ International Conference on Multisensor Fusion and Integration for Intelligent Systems* (IEEE, Piscataway, NJ, 1996), pp. 727–735.
[13] J. A. Simmons and L. Chen, ''The acoustic basis for target discrimination by fm echo locating bats,'' J. Acoust. Soc. Am. **86**, 1333–1350 (1989).

The acoustic finite integration technique for waves of cylindrical symmetry (CAFIT)

A. Peiffer[a] and B. Köhler[b]
Fraunhofer-Institute for Nondestructive Testing, Branchlab Dresden EADQ, Krügerstr. 22, 01326 Dresden, Germany

S. Petzold[c]
Institute for Biomedical Engineering, Technical University, Dresden, Germany

(Received 26 September 1996; revised 7 April 1997; accepted 23 April 1997)

Many ultrasonic nondestructive testing applications have cylindrical geometries. Examples involve the excitation of ultrasound by cylindrical piezoelectic probes or by laser, x rays, electron beams [A. C. Tam, Rev. Mod. Phys. **58**, 381–431 (1986)], or ion beams [L. Sulak *et al.*, Nucl. Instrum. Methods **161**, 203–217 (1979)]. Thus, calculations of cylindrical wave propagation are important for a better understanding and interpretation of many testing situations. This paper deals with the AFIT Code or finite volume method for numerical simulation of sound propagation in fluids adapted to cylindrical geometries (CAFIT). A comparison is made with standard difference-equations techniques also utilized for cylindrical geometries. Two examples are dealt with: (1) The sound generation by a high energy beam of heavy ions stopping in water; (2) the multimode sound propagation in a medical doppler injection device excited by a disk probe. © *1997 Acoustical Society of America.* [S0001-4966(97)03208-6]

PACS numbers: 43.20.Bi, 43.20.Mv, 43.35.Yb, 43.80.Vj [JEG]

INTRODUCTION

The finite volume method for acoustical wave propagation or acoustic finite integration technique (AFIT) is a very stable tool for the numerical calculation of wave propagation in fluids.[1,2] Until now, this code has been used for calculations in the Cartesian coordinate system.

Solving real three-dimensional problems in a three-dimensional grid usually exceeds the capacities and capabilities of any standard workstation available today. Thus, problems need to be reduced to two dimensions, by assuming the wave propagation translational invariant in one dimension. This means for example, replacing a rectangular transducer by a linear strip. But this is always an approximation of the real problem.

For problems of cylindrical geometry, the wave motion can be described exactly by a wave equation only in the two coordinates z and ϱ. Therefore, this paper aims to set up a new discrete scheme for numerical simulation of wave propagation in this system based on the AFIT ideas and using only a two-dimensional grid. Unfortunately, it is not possible without special extentions of this algorithm to deal with the tangential component.

Cylindrical systems have a unique property for performing numerical calculations i.e., the singularity along the z axis. This singularitiy can be easily managed by using our proposed scheme, because of the involvement of volume-integral formulations.

Two fields of applications will be discussed in this paper: First, the sound generation by a beam of high energy particles (ion-acoustic source) and second, the sound field of a disk probe used in a doppler injection device. The latter example is a good illustration for multimode wave propagation in cylindrical systems.

I. BASIC EQUATIONS

The basic equations are the linear equation of continuity

$$\frac{\partial \rho^*}{\partial t^*} + \rho_0^* \nabla^* \mathbf{v}^* = \dot{M}^*, \tag{1}$$

and the linear equation of motion

$$\rho_0^* \frac{\partial \mathbf{v}^*}{\partial t^*} = -\nabla^* p^* + \mathbf{f}^*. \tag{2}$$

The asterisk symbol above the variable indicates that it has a dimension. Variables without an asterisk symbol are dimensionless. Here $\mathbf{v}^*$, $\mathbf{x}^*$, $\mathbf{f}^*$ are the Cartesian vectors of velocity, space, force density and ρ^*, ρ_0^* are the density and equilibrium density, respectively. The acoustic pressure is p^*, and M^* a density source.

The introduction of the following dimensionless quantities

$$p = \frac{p^*}{\rho_0^* c_0^{*2}} \quad t = \frac{t^* c_0^*}{l_0^*} \quad \mathbf{x} = \frac{\mathbf{x}^*}{l_0^*}, \tag{3}$$

$$\mathbf{v} = \frac{\mathbf{v}^*}{c_0^*} \quad M = \frac{M^*}{\rho_0^*} \quad \mathbf{f} = \frac{\mathbf{f}^*}{\rho_0^* c_0^{*2} l_0^*}, \tag{4}$$

and the isentropic speed of sound

$$c_0^{*2} = \frac{\partial p^*}{\partial \rho^*}$$

[a]Electronic mail: peiffer@eadq.izfp.fhg.de
[b]Electronic mail: koehler@eadq.izfp.fhg.de
[c]Electronic mail: petzold@rcs.urz.tu-dresden.de

leads to dimensionless equations which are more useful for numerical calculation, because pressure and velocity have the same orders of magnitude:

$$\frac{\partial p}{\partial t} = -\nabla \mathbf{v} + \dot{M}, \tag{5}$$

$$\frac{\partial \mathbf{v}}{\partial t} = -\nabla p + \mathbf{f}. \tag{6}$$

II. RECEIVING THE DIFFERENCE EQUATIONS

The main benefit of the AFIT procedure is, that the *staggered grid* leads to a better accuracy, at much lower efforts in programming and calculation time. This simple central difference approximation gives better results than second-order schemes based on nonstaggered grids. The key procedure of AFIT consists of performing a control volume integration for Eqs. (5) and (6) and transforming every integral over the gradient or divergence terms into a surface integral by using the Gauss law. This leads automatically to a staggered grid. In the two-dimensional case the volume cells correspond to an area cell and the volume faces to the borders of the area.

There are two possibilities to obtain the discrete equations in cylindrical coordinates:

(1) Taking the cylindrical differential equations and using a ''Cartesian'' two-dimensional grid for evaluation of the curve and area integration in radial and axial coordinates ϱ and z.
(2) Using the equations in Cartesian form and performing the integration over control volumes of cylindrical shape.

Both methods give the same result. As the second is more instructive the integration is carried out according to this method. A further advantage results from the better treatment of the singular term along the z axis.

Though only wave propagation in two coordinates is described, it has to be dealt with three-dimensional control volumes, assuming the tangential component to be zero. This follows from the fact that the volume faces perpendicular to the tangential component contribute to the radial component.

A. Conventions

In order to achieve a good description of cylindrical wave propagation, a well suited control volume has to be chosen, that is, a ''*piece of cake shaped*'' sector volume. The Cartesian and cylindrical unit vectors and coordinates are also shown. v_x, v_y and v_z are the Cartesian components of the velocity $\mathbf{v}$, whereas u, v, and w are the radial, tangential and axial component in terms of a cylindrical system (Fig. 1). Setting v to zero we have

$$\mathbf{v} = v_x \hat{\mathbf{e}}_x + v_y \hat{\mathbf{e}}_y + v_z \hat{\mathbf{e}}_z = u \hat{\mathbf{e}}_\varrho + w \hat{\mathbf{e}}_z .$$

Though the pressure and velocity field is cylindrical, the volume integration is performed in a Cartesian coordinate system. Faces and control volume are given by cylindrical parameters ϱ, φ and z. Thus, the Cartesian components of the velocity are

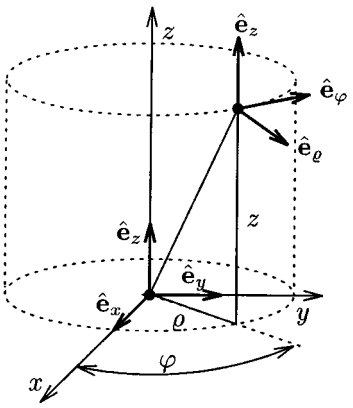


FIG. 1. Coordinate system conventions.

$$v_x(\varrho,\varphi,z) = u(\varrho,z)\cos(\varphi),$$

$$v_y(\varrho,\varphi,z) = u(\varrho,z)\sin(\varphi),$$

$$v_z(\varrho,\varphi,z) = w(\varrho,z).$$

Each control volume, introduced in the following, is localized around a center gridpoint denoted by the upper index (c). The top and bottom volume faces are denoted by the indices (t) and (b), the inner and outer cylinder sector walls by (i) and (o) and the side faces by $(+)$ and $(-)$ according to a positive or negative angle.

B. Equation of continuity

The volume integral of Eq. (5) leads to:

$$\int_{V_p} \frac{\partial p}{\partial t} dV = -\int_{V_p} \nabla \mathbf{v} dV + \int_{V_p} \dot{M}$$

$$= -\oint_{\partial V_p} \mathbf{v} d\mathbf{s} + \int_{V_p} \dot{M} dV. \tag{7}$$

Figure 2 shows the control volume V_p for the pressure. The mean pressure is represented by the grid point in the center of the cell. Each grid point of the velocity component perpendicular to the faces is located at the face center.

Only the velocity components perpendicular to the surface contributes to the first term on the RHS of Eq. (7). These are u for the inner and outer one and w for the top and bottom faces. The sides need not be considered, as the tangential component v perpendicular to these faces is zero. Evaluation of the integrals gives

$$\int_{V_p} \frac{\partial p}{\partial t} dV = \dot{p}^{(c)} \varrho \Delta\varrho \Delta z \Delta\varphi, \tag{8}$$

$$\oint_{\partial V_p} \mathbf{v}\, d\mathbf{s} \cong +\left[\left(\varrho + \frac{\Delta\varrho}{2}\right) u^{(o)} - \left(\varrho - \frac{\Delta\varrho}{2}\right) u^{(i)}\right] \Delta z \Delta\varphi$$

$$+ [w^{(t)} - w^{(b)}] \varrho \Delta\varrho \Delta\varphi. \tag{9}$$

The integral over $\dot{M}$ is treated similar. Dividing by the cell-volume $\Delta V = \varrho \Delta\varrho \Delta\varphi \Delta z$ one gets

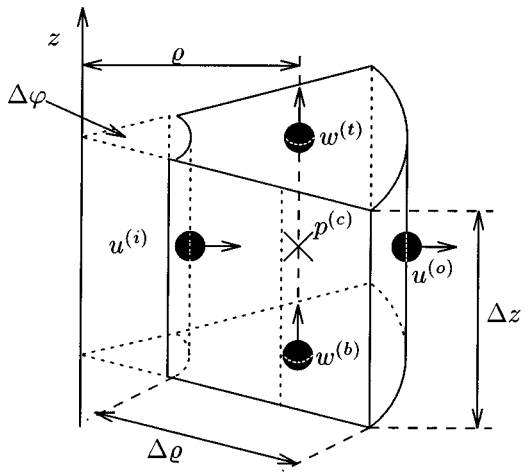


FIG. 2. Control volume for p.

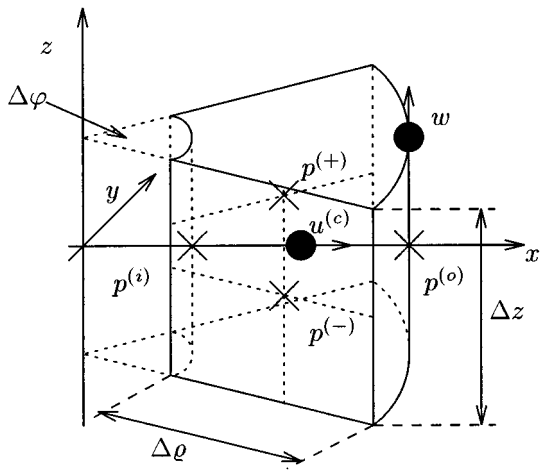


FIG. 3. Control volume for u.

$$\dot{p}^{(c)} = -\frac{u^{(o)}-u^{(i)}}{\Delta\varrho} - \frac{w^{(t)}-w^{(b)}}{\Delta z} - \frac{u^{(o)}+u^{(i)}}{2\varrho} + \dot{M}^{(c)}. \tag{10}$$

Here the scalar quantity p is averaged. Thus, the shape of the volume cell is not essential for consideration. For better interpretation, the equation of continuity in terms of cylindrical coordinates is given as follows:

$$\frac{\partial p}{\partial t} = -\frac{\partial u}{\partial \varrho} - \frac{\partial w}{\partial z} - \frac{u}{\varrho} + \dot{M}. \tag{11}$$

The third term on the RHS of Eq. (10) may be interpreted as the approximation of u/ϱ in Eq. (11), because it represents the linear interpolation of u/ϱ at the center of the control volume. Thus, only the choice of the well fitted control volume leads to discrete formulation of the cylindrical differential equation. Not only Eq. (11) but also Eq. (10) is singular for $\varrho = 0$. Thus, the z axis must be dealt with separately.

C. Equation of motion

In the case of Eq. (6), the volume integration is performed for a vector quantity. Surface integrals arise over scalar quantities, like the pressure p:

$$\int_V \frac{\partial \mathbf{v}}{\partial t} dV = -\int_V \nabla p \; dV + \int_V \mathbf{f} \; dV$$

$$= -\oint_{\partial V} p \; d\mathbf{s} + \int_V \mathbf{f} \; dV. \tag{12}$$

The mean value of every velocity component is not coupled for Cartesian coordinate systems, but averaging over every radial velocity component u the value must be projected to the direction of the centered vector $u^{(c)}$.

The cylindrical sector-shaped volume V is placed in the Cartesian coordinate system, so that the cylindrical velocity components u and w at the center of each volume cell coincide with the Cartesian coordinate v_x and v_z.

Evaluating each component v_x and v_z in Eq. (12)—taking the cylindrical symmetry of the considered solutions into account—leads automatically to the difference equations for the velocity grid points. The volume integration gives

$$\int_{V_u} \frac{\partial v_x(\varrho,\varphi,z)}{\partial t} dV = \int_{V_u} \dot{u}(\varrho,z)\cos(\varphi) dV$$

$$= \dot{u}^{(c)} 2\varrho \Delta\varrho \Delta z \sin\frac{\Delta\varphi}{2}, \tag{13}$$

$$\int_{V_w} \frac{\partial v_z(\varrho,\varphi,z)}{\partial t} dV = \dot{w}^{(c)} \varrho \Delta\varrho \Delta z \Delta\varphi, \tag{14}$$

$$\int_{V_v} \frac{\partial v_y(\varrho,\varphi,z)}{\partial t} dV = 0. \tag{15}$$

The components of the force density $\mathbf{f}$ are also similarly treated. It should be emphasized that every component has a different control volume, because each velocity component node is localized at a different place as shown in Figs. 2 and 3. Obviously the y component in the tangential direction vanishes.

The parameter expressions of the face element vectors $d\mathbf{s}_{(\cdot)}$ of the cell volume are given by

$$d\mathbf{s}_{(o/i)} = \pm\left(\varrho \pm \frac{\Delta\varrho}{2}\right)\begin{pmatrix}\cos\varphi'\\ \sin\varphi'\\ 0\end{pmatrix} d\varphi' \; dz', \tag{16}$$

$$d\mathbf{s}_{(+/-)} = \begin{pmatrix}-\sin\frac{\Delta\varphi}{2}\\ \pm\cos\frac{\Delta\varphi}{2}\\ 0\end{pmatrix} d\varrho' \; dz', \tag{17}$$

$$d\mathbf{s}_{(t/b)} = \pm\begin{pmatrix}0\\ 0\\ \varrho'\end{pmatrix} d\varrho' \; d\varphi', \tag{18}$$

with parameter ranges:

$$\varrho' \in \left[\varrho - \frac{\Delta\varrho}{2}, \varrho + \frac{\Delta\varrho}{2}\right],$$

$$z' \in \left[z - \frac{\Delta z}{2}, z + \frac{\Delta z}{2}\right], \quad \varphi' \in \left[-\frac{\Delta\varphi}{2}, \frac{\Delta\varphi}{2}\right].$$

Using this parameter forms we arrive at the surface integrals of the scalar quantity p:

$$\int_{\partial V_{(o/i)}} p \; d\mathbf{s}_{(o/i)} \cong p^{(o/i)} \int_{\partial V_{(o/i)}} d\mathbf{s}_{(o/i)} = \pm 2 p^{(o/i)} \Delta z \left(\varrho \pm \frac{\Delta\varrho}{2}\right) \sin\frac{\Delta\varphi}{2}\, \hat{\mathbf{e}}_x, \tag{19}$$

$$\int_{\partial V_{(+/-)}} p \; d\mathbf{s}_{(+/-)} \cong p^{(+/-)} \Delta z \Delta\varrho = \left(\sin\frac{\Delta\varphi}{2}\, \hat{\mathbf{e}}_x \mp \cos\frac{\Delta\varphi}{2}\, \hat{\mathbf{e}}_y\right), \tag{20}$$

$$\int_{\partial V_{(t/b)}} p \; d\mathbf{s}_{(t/b)} \cong p^{(t/b)} \varrho \Delta\varrho \Delta\varphi \; \hat{\mathbf{e}}_z. \tag{21}$$

The volume faces $\partial V_{(o/i)}$ and $\partial V_{(t/b)}$ contributes to u or w, but $\partial V_{(+/-)}$ influences u and—if regarded—v. Here it can be seen that for a correct treatment, the points $p^{(+)/(-)}$ are required for u. Thus, $p^{(+)/(-)}$ must be obtained by linear interpolation from $p^{(o)}$ and $p^{(i)}$. The need for this interpolation is not a disadvantage of the two-dimensional grid. Taking a three-dimensional axially symmetrical grid would result in an interpolation from four gridpoints. This leads to numerical problems which cannot be discussed here and are one reason for not considering solutions having a tangential component.

Considering that $\hat{\mathbf{e}}_x$ and $\hat{\mathbf{e}}_z$ are parallel to the component $u^{(c)}$ and $w^{(c)}$ and by taking Eqs. (10), (13), (14) and (19)–(21) one finally gets the complete set of difference equations:

$$\dot{u}^{(c)} = -\frac{p^{(o)} - p^{(i)}}{\Delta\varrho} + f_\varrho^{(c)}, \tag{22}$$

$$\dot{w}^{(c)} = -\frac{p^{(t)} - p^{(b)}}{\Delta z} + f_z^{(c)}, \tag{23}$$

$$\dot{p}^{(c)} = -\frac{u^{(o)} - u^{(i)}}{\Delta\varrho} - \frac{w^{(t)} - w^{(b)}}{\Delta z} - \frac{u^{(o)} + u^{(i)}}{2\varrho} + \dot{M}^{(c)}. \tag{24}$$

Here it can be seen again that the cylindrical differential equations have the same structure:

$$\frac{\partial u}{\partial t} = -\frac{\partial p}{\partial \varrho} + f_\varrho \quad \frac{\partial w}{\partial t} = -\frac{\partial p}{\partial z} + f_z. \tag{25}$$

Using the discrete equations the propagation of sound is received by calculating *marching-in-time* the pressure and velocity components in the numerical grid. As explained by Marklein in Ref. 1 the time-discrete steps for velocity and pressure are also staggerd. That means for every grid point p_ξ, u_ξ, or w_ξ

$$p_\xi(t) \to p_\xi((n - \tfrac{1}{2})\Delta t) = p_\xi[n - \tfrac{1}{2}],$$

$$u_\xi(t) \to u_\xi(n\Delta t) = u_\xi[n],$$

$$w_\xi(t) \to w_\xi(n\Delta t) = w_\xi[n].$$

Each new value is calculated from the value of the last time step and the time derivative in the middle of the time interval, which is obtained by the above difference Eqs. (22), (23), and (24):

$$p_\xi[n + \tfrac{1}{2}] = p_\xi[n - \tfrac{1}{2}] + \Delta t \;\; \dot{p}_\xi[n],$$

$$u_\xi[n+1] = u_\xi[n] + \Delta t \;\; \dot{u}_\xi[n + \tfrac{1}{2}],$$

$$w_\xi[n+1] = w_\xi[n] + \Delta t \;\; \dot{w}_\xi[n + \tfrac{1}{2}] \quad n = 0,1, \ldots .$$

The grid constants $\Delta\varrho$ and Δz depend on the smallest wavelength that occurs in the according direction of the sound field. In order to achieve good accuracy we chose $\Delta\varrho = \lambda_\varrho/15$ and $\Delta z = \lambda_z/15$. λ_ϱ and λ_z are the smallest wavelengths in the ϱ or z direction. The Courant criterion requires that the time step Δt be chosen such that

$$\Delta t \leqslant \frac{1}{\sqrt{1/\Delta r^2 + 1/\Delta z^2}}.$$

The least numerical dispersion is obtained right at the Courant limit for Cartesian grids. A detailed von Neumann analysis of the cylindrical problem is omitted here. As the Courant criterion is only necessary and may not be sufficient, we chose Δt by trial. In Ref. 1 it is said that numerical stability is always assured by choosing $\Delta t = 0.5$ times the smallest grid distance for Cartesian grids. We can confirm this also for the cylindrical grid.

D. The axis

The control volume at the axis is a cylindrical-sector volume touching it (Fig. 4). The first idea might be to set the p-grid points at the axis in order to simply avoid $\varrho = 0$ in Eq. (10). But there is no inner volume face of the control volume near the axis. Due to the above fact, there must not be a considerable grid point. If the pressure points are along the axis the CAFIT procedure cannot deal with those values. Therefore, the values of the grid point must vanish along the axis. This is only possible for the radial velocity component u, which—according to Eq. (11)—must be zero and that is not possible for the pressure which definitely has values not equal to zero.

The difference equation due to the grid point Δz beside the axis is:

$$\dot{p}^{(c)} = -\frac{2u^{(o)}}{\Delta\varrho} - \frac{w^{(t)} - w^{(b)}}{\Delta z}. \tag{26}$$

Setting the velocity to zero indicates that the axis behaves like a small rigid cylinder or wire.

The control volume of the first radial velocity grid point does not touch the axis, thus it can be calculated from the normal difference equation (24).

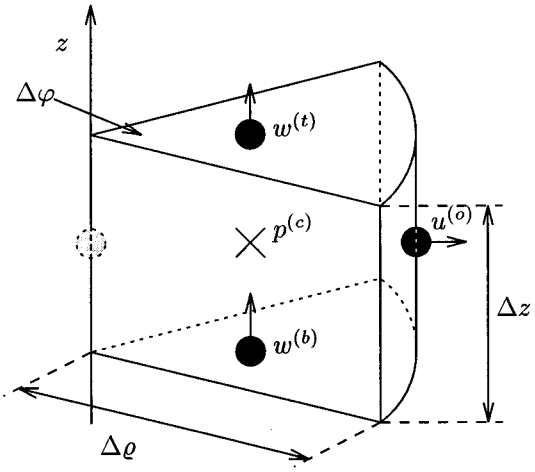


FIG. 4. Volume cell at the axis.

It should be noted that for other numerical difference methods without a staggered grid, it is much more complicated to guarantee $u^{(a)}$ to be zero, because a radial pressure gradient has to be avoided along the axis as this would lead to not vanishing values for the radial velocity u.[3]

E. The grid

Finally, the chosen grid performance is shown in Fig. 5. The radial velocity points are assumed to be zero at the axis and are not represented in the grid arrangement. Boundaries are easy to set up by positioning a pressure (or velocity) gridpoint exactly at the boundary and arranging the velocity (or pressure) points before and behind it, for describing the change of impedance there.

III. TEST CASES

In order to prove the accuracy and applicability of the AFIT algorithm for cylindrical geometries, two numerical test cases are set up. Ottitsch[3] uses the reflection of a spherical wave into itself for evaluation of several finite difference schemes. This is a very critical case, as nearly the whole energy is conserved in few grid points during the ''reflection'' at the center. Thus, this test case is also used here to test the CAFIT code. A second test case of more practical interest is the ratiation of sound by a cylindrical piston in a baffle.

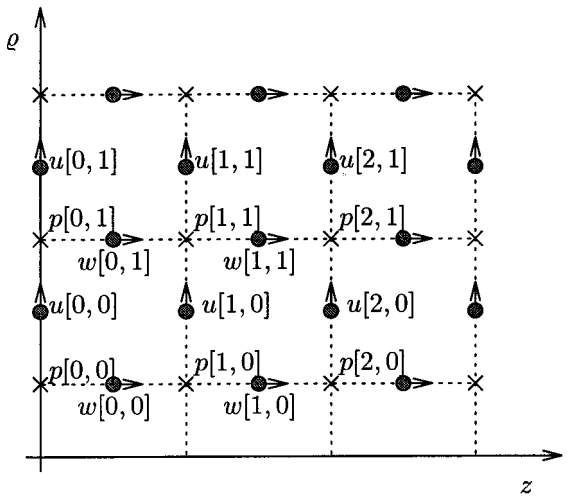


FIG. 5. The grid in the cylindrical half plane.

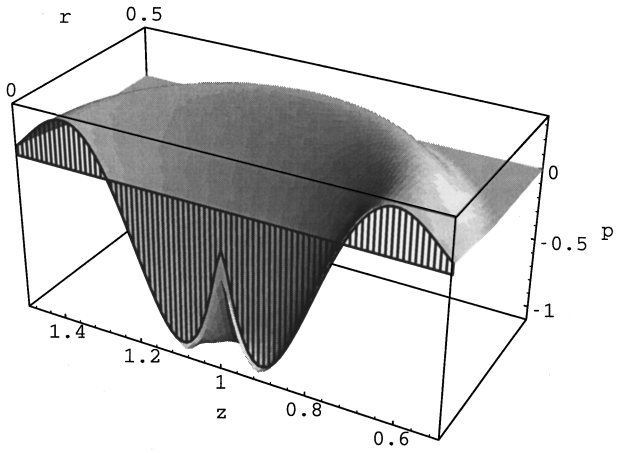


FIG. 6. Sound field, $t=-0.05$.

A. Spherical wave test case

Ottitsch choses the following initial conditions for p and v_l, the spherical velocity pointing away from the center:

$$p=-\frac{3.84(l+t)^2}{\epsilon^6 l}*$$
$$[\epsilon^3-4\epsilon^2(l+t)+5\epsilon(l+t)^2-2(l+t)^3],$$

$$v_l=\frac{3.2(l+t)^2}{\epsilon^6 l^2}*[\epsilon^3(2l-t)-3\epsilon^2(3l^2+2lt-t^2)$$
$$+3l(4l^3+7l^2+2lt^2-t^3)-5l^4$$
$$-t(14l^3+12l^2t+2lt^2-t^3)],$$

$$l=\sqrt{r^2+(z-1)^2},$$

with $\epsilon=0.55$, $t=-0.4$, and $0\leqslant l+t<\epsilon$. In all calculations the radial coordinate is named r instead of ϱ. The numerical grid of 160×80 points represents the half plane with $z=0\cdots2$ and $r=0\cdots1$. The time step is $\Delta t=0.7\Delta r$ and $\Delta r=\Delta z$, l is the distance from the center, which is at $z=1$ and $r=0$. The analytical solution of such a wave can easily be calculated.

Figures 6 and 7 show the wave for two different time steps calculated by CAFIT. The grid in front of both figures shows the analytical solution. The result is stable; only some slight numerical oscillations can be noticed at the back of the wave. In Ottitsch's work,[3] only the MacCormack scheme gives a comparable result.

Figure 8 shows comparison of the pressure obtained by the analytical solution and the numerical results of CAFIT and the MacCormack schemes. The pressure values of CAFIT are taken from a line perpendicular to the axis; those from the MacCormack scheme along the axis. The results of the MacCormack scheme differ more, in spite of the higher

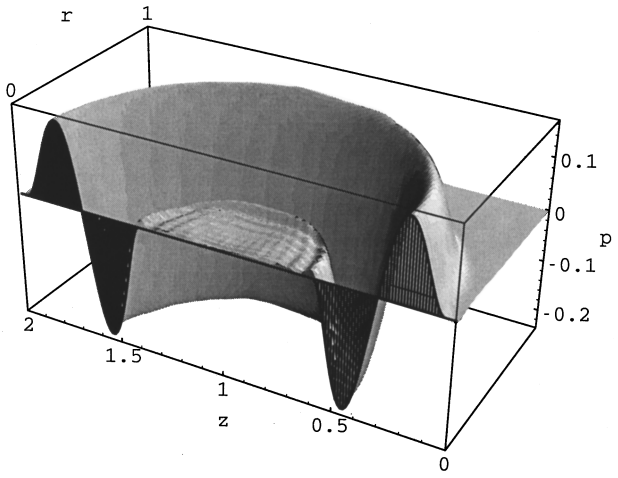


FIG. 7. Sound field, $t=0.95$.

efforts in programming this two step algorithm.[3] The grid has been chosen rather roughly, in order to make a direct comparison to Ottitsch's work[3] possible. A finer grid (320 $\times$640) gives very good agreement between numerical and analytical solution (Fig. 9).

This test case proves that this numerical code is capable of calculating cylindrical wave propagation, especially the reflection at the axis (which is a superposition of passing waves in reality).

B. Piston test case

The second test deals with the transient radiation of a cylindrical piston of radius a, set in the rigid boundary plane at $z=0$ and driven with velocity $w_p(t)$. An analytic expression of the transient pressure field is cited by Harris.[4] The pressure is given by two convolution integrals valid for regions I or II:

$$p(z,r,t)=\frac{1}{\pi}\int_0^{\theta_1} w_p(t-R_1(\theta))-w_p(t-R_2(\theta))d\theta$$

$$\text{I: } r>a, \tag{27}$$

$$p(z,r,t)=w_p(t-z)-\frac{1}{\pi}\int_0^{\pi} w_p(t-R_2(\theta))d\theta \quad \text{II: } r\leqslant a. \tag{28}$$

The limit of integration is $\theta_1=\arcsin(a/r)$. $R_1(\theta)$ and $R_2(\theta)$ are the distances between the point of observation at (z,r) and the ''front'' and ''rear'' edge of the piston. θ is the parameter of integration:

$$R_{1/2}(\theta)=\sqrt{z^2+(r\cos\theta\mp\sqrt{a^2-r^2\sin^2\theta})^2}.$$

Only for points at the axis ($r=0$) may Eq. (28) be written in a closed form:

$$p(z,0,t)=w_p(t-z)-w_p(t-\sqrt{a^2+z^2}). \tag{29}$$

The piston is driven by the velocity:

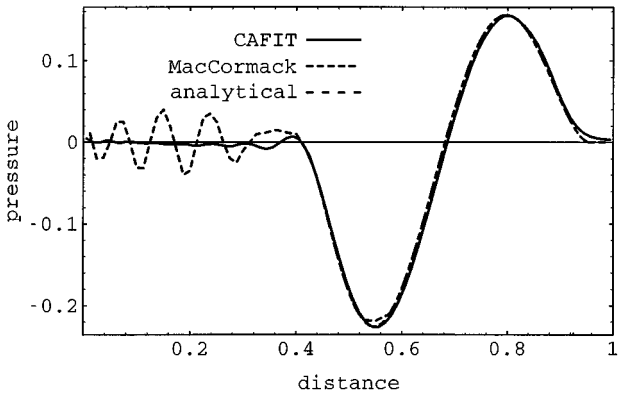


FIG. 8. Numerical and analytical solution (80$\times$160).

$$w_p(t)=\begin{cases}\sin^2(t*\pi), & 0\leqslant t\leqslant 1\\ 0, & \text{otherwise.}\end{cases}$$

For showing the limit of this code we chose a very rough grid with $\Delta r=\Delta z=0.1$. We take a finer grid with $\Delta r=\Delta z=0.025$ in order to prove that a very good agreement between the analytical and numerical solution can be achieved. For both grids $\Delta t=0.7\Delta z$. Using 60$\times$40 points or 240$\times$160 points, respectively, the grids represent the half plane with $z=0\cdots 6$ and $r=0\cdots 4$. Assuming a bandwith of the exciting signal $w_n(t)$, defined by the spectral amplitude of 10 percent of the maximum, the smallest wavelength is $\lambda_{\min}=0.625$. Note that $\Delta r=\Delta z$ is approximately $\lambda_{\min}/6$ for the coarse grid.

As the axis is supposed to be the critical region for the numerical scheme, the analytical and numerical time signals at $r=0$ and $z=1$, 3, 7 are compared in cases (a)–(c) in Figs. 10 and 11. The evaluation of Eqs. (27) and (28) for $r\neq 0$ must be performed by numerical integration. This is done for $z=2$ and $r=2$ in case (d).

In Fig. 10 the results for the coarse grid are shown. Good agreement can be seen between both types of solutions. That the finer grid leads to a very good correspondence is shown in Fig. 11. In Fig. 12 a snapshot of the sound field at $t=3$ is shown for the coarse grid. Some numerical oscillations occur behind the pulse. The semi-circle in front denotes the position of the piston.

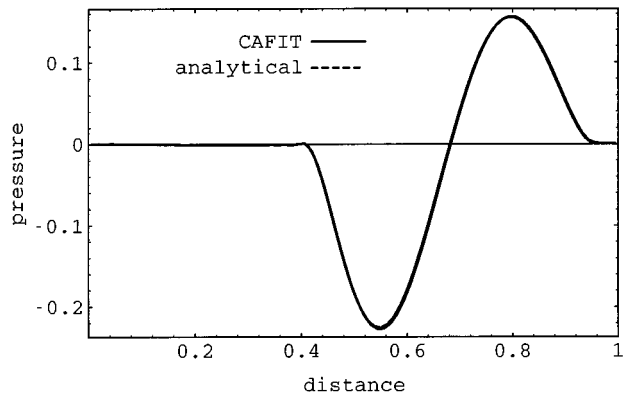


FIG. 9. Numerical and analytical solution (320$\times$640).

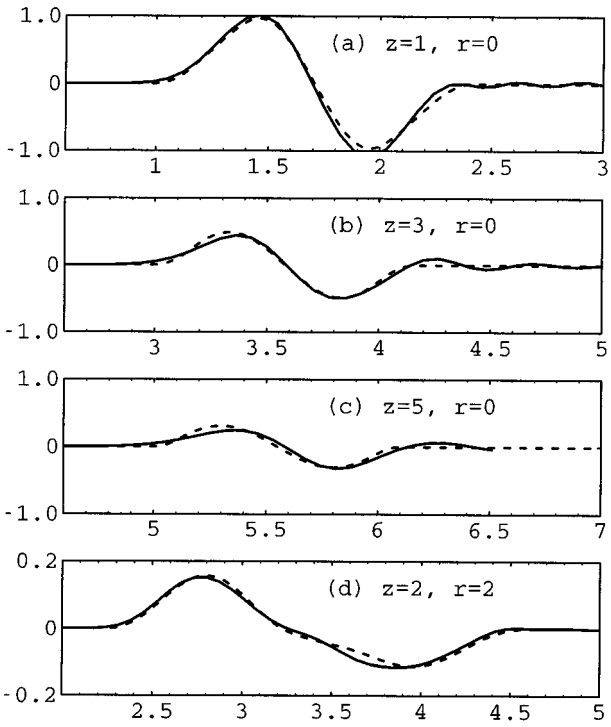


FIG. 10. Numerical and analytical time signals for coarse grid ($\Delta r=\Delta z=0.1$) at several positions (a)–(d). The solid line represents the numerical solutions, the dashed one the analytical.

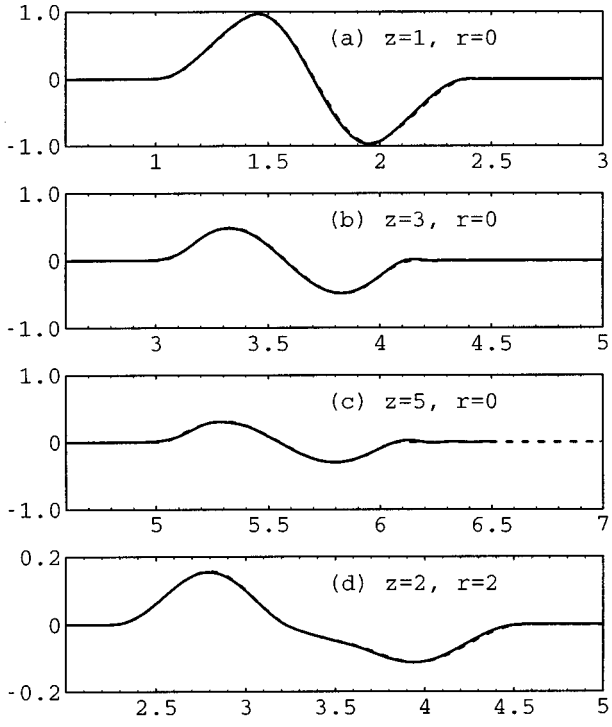


FIG. 11. Numerical and analytical time signals for fine grid ($\Delta r=\Delta z=0.025$) at several positions (a)–(d). The solid line represents the numerical solutions, the dashed one the analytical.

C. Test case conclusions

Both cases show that the CAFIT scheme is capable of calculating cylindrical wave propagation. Not only for the reflection at the axis, but also for the radiation from cylindrical sources is this code consistent and stable.

IV. EXAMPLES

A. Ion-acoustic source

A short-pulsed heavy-ion beam of high energy forms a thermoelastic source that is well described by the initial condition:[5]

$$p(z,r,t=0)=\frac{\beta^* c_0^{*2}}{c_p^*}\Psi(r,z)\frac{dE}{dz}$$

$$=\frac{\Gamma}{2\pi\sigma^2}\exp\left(-\frac{r^2}{2\sigma^2}\right)\frac{dE}{dz},$$

$$u(z,r,t=0)=w(z,r,t=0)=0,$$

where c_p^* is the specific heat at constant pressure, β^* the thermal expansion coefficient and Γ the so-called Grüneisen parameter, which is a measure for the efficiency of the thermoelastic sound generation process. Ψ is the function of the radial energy density, σ the Gauss radius of the beam profile and dE/dz of the distribution along the axis, the so-called linear energy transfer (LET). There is a maximum at the end of the beam track called the Bragg peak. The computation is made due to the data of a beam of carbon ions (^{12}C) of energy 300 MeV/u. The length l_0^* is 1 cm. The pressure distribution of the initial condition ($t=0$) is shown in Fig. 13.

The front boundary of the medium where the beam enters is assumed to be rigid. In Figs. 14 and 15 the sound fields at the time $t=0.3$ and $t=2$ are shown. Only the region near the Bragg peak is shown, as the most interesting aspects are noticed here. The radiation of sound from the thermoelastic source is clearly visible. The propagation of the spherical wave due to the Bragg peak and the cylindrical wave due to the constant LET along the beam track can also be noticed in these figures.

In order to perform a further test, these calculations are made for a time signal at position $z=20$, $r=0$. In Fig. 16 the

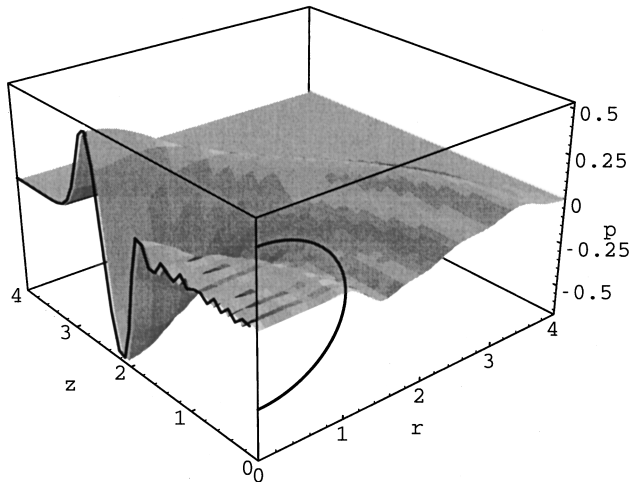


FIG. 12. Sound field of piston in a baffle, $t=3.0$, coarse grid (60×40).

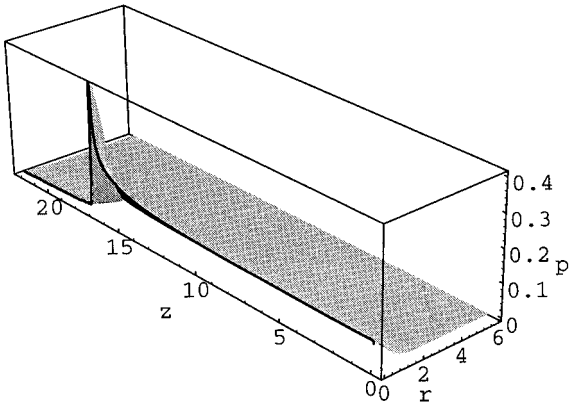


FIG. 13. Initial pressure condition ($t=0$).

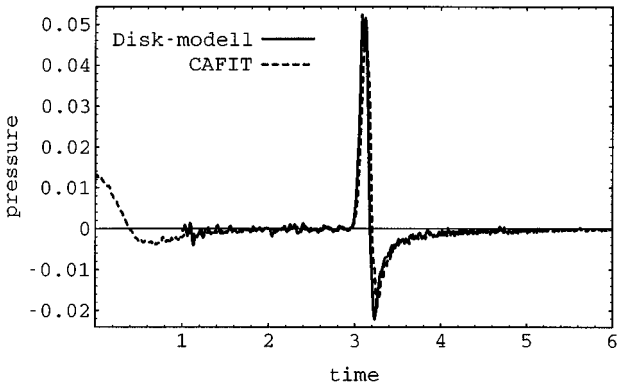


FIG. 16. Comparison of CAFIT and disk model.

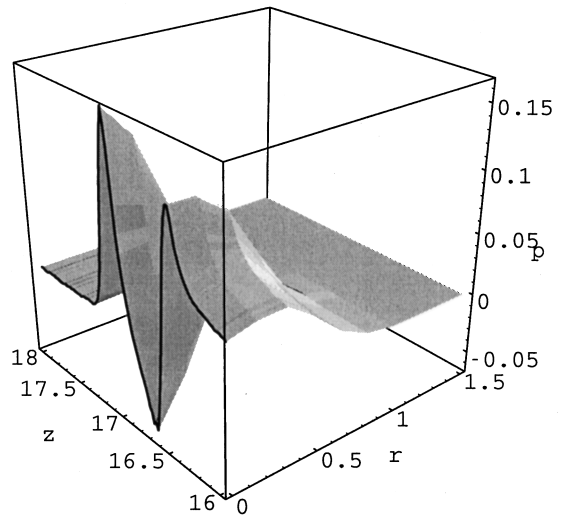


FIG. 14. Detailed pressure field ($t=0.3$).

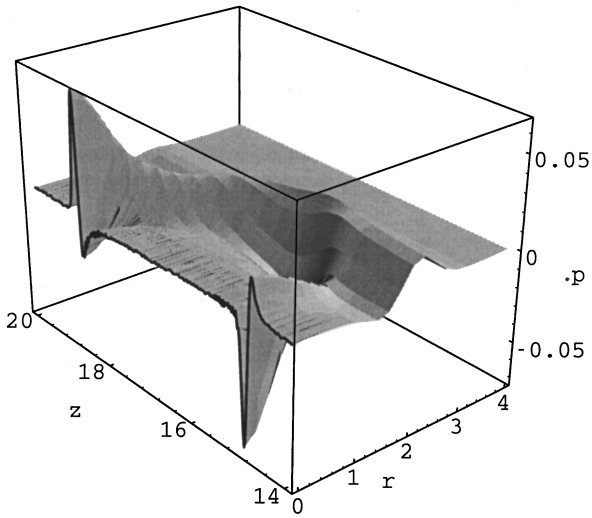


FIG. 15. Detailed pressure field ($t=2.0$).

results are compared to the calculation of the same problem performed by the superposition of gaussian disk sources.[6]

B. Doppler injection device

Development of the Doppler injection device in the field of biomedical engineering is useful to achieve more safety for punctures of arteries and veins. An ultrasonic Doppler transducer (circular piston transducer) is fixed to the plunger of a liquid filled injection device. The problem is to guide the ultrasonic waves through the injection into the needle with high efficiency. Due to the cylindrical geometry of injection devices, the CAFIT procedure is appropriate to calculate this guided wave propagation and the manifold reflections at the walls of injection and needle.

In Fig. 17 the geometry of the modelled injection is shown. In order to excite many higher order modes for evident presentation, a flat angle is chosen for the cone adapter.

The center frequency of excitation is $f_0^*=4$ MHz. The unit length is $l_0^*=1$ mm, $c_0^*=1500$ m/s. Every boundary is assumed to be rigid, as well as the plunger and the piston. The radiation of the piston transducer is modelled by setting the velocity w_p uniformly at the piston surface to:

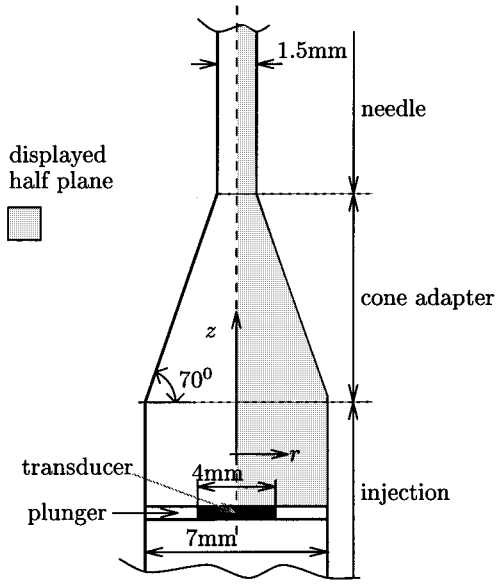


FIG. 17. Geometry of Doppler injection device.

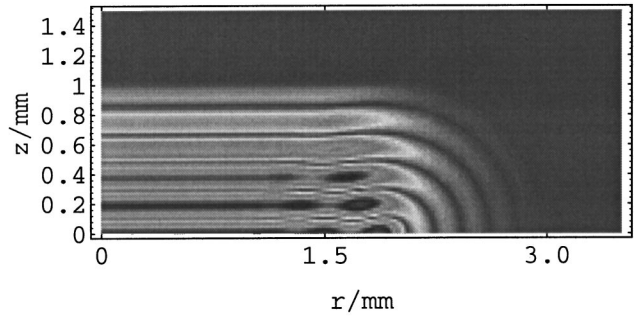


FIG. 18. Wave fronts near the transducer ($t=1.032$).

$$w_p(t)=\begin{cases}\sin(2\pi f_0 t)\sin\left(\dfrac{2\pi f_0 t}{10}\right), & 0\leqslant t\leqslant \dfrac{5}{f_0}\\ 0, \quad \text{otherwise.}\end{cases}$$

In the following, the simulation results are shown by snapshots of the pressure magnitude at different times.

In Fig. 18, the sound wave at time $t=1.032$ is shown. The plane wave and the edge wave component coming from the radiators rim can be observed.

Figure 19 shows the sound wave travelling through the adapter into the needle. The parts of the wave fronts near the axis enter the needle nearly undisturbed, whereas the outer parts—multiplied at the wall reflected—enter needle and adapter later and at different angles. This is the reason for the excitation of higher mode sound propagation in the needle.

In order to identify the mode order one has to count the maxima along the cross section of the needle. The higher the mode, the slower it propagates along the axis. Thus, the distance between the modes increases during their passing through the needle (Fig. 20).

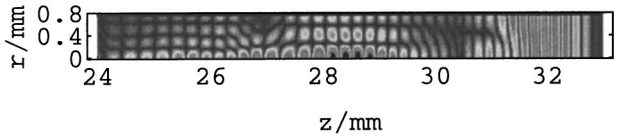


FIG. 20. Higher order mode propagation in the needle.

V. CONCLUSIONS

The CAFIT code provides a suitable tool for calculating cylindrical wave propagation. It may be utilised for the computation of many probe characteristics in nondestructive ultrasonic testing applications, using cylindrical probe transducers in fluids. Further applications are the description of arbitrary thermoacoustic sources, like laser, ion, or electron beams, which mostly have a cylindrical geometry. Sound propagation in pipes may be calculated, for instance to support the construction of sound absorbers. For this purpose the CAFIT code has to be extended to the Navier–Stokes equations including dissipative terms.

This cylindrical code is also extended to elastodynamic equations for describing sound propagation in solids. Nevertheless, this code can be used for initial evaluation of the pressure waves, neglecting the shear wave components.

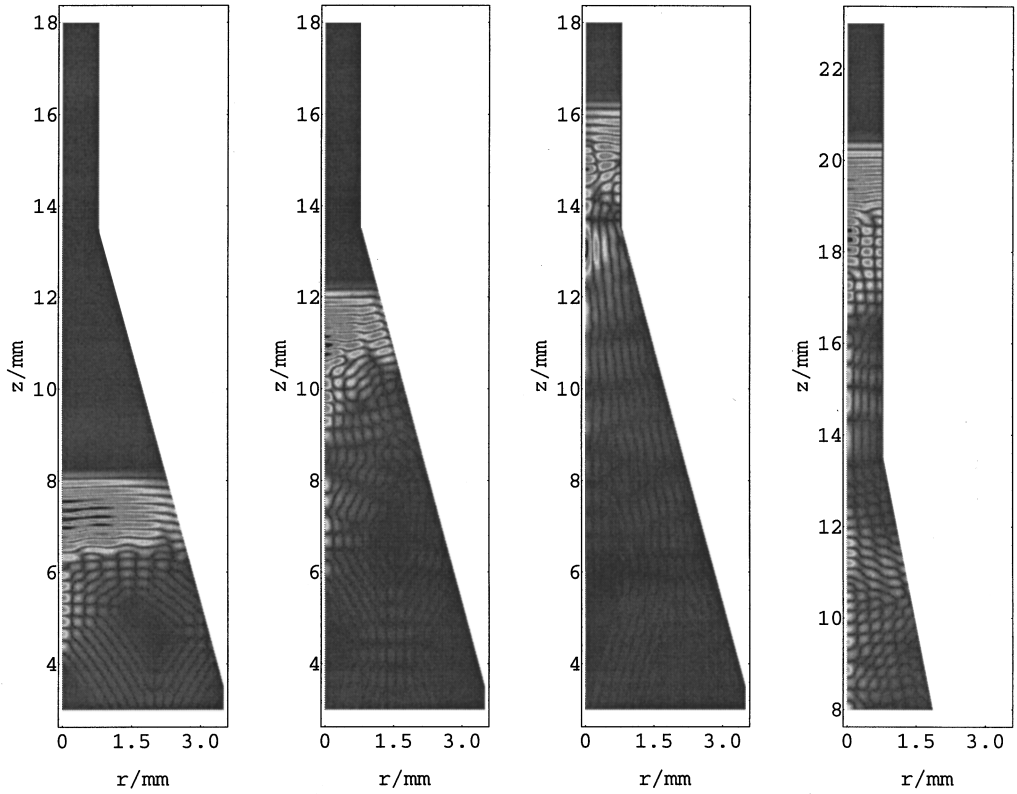


FIG. 19. Wave entering the needle; time steps $t=8.25$, $t=12.375$, $t=16.5$ and $t=20$.

ACKNOWLEDGMENTS

Support for this research was provided in part by the Gesellschaft für Schwerionenforschung (GSI) in Darmstadt. The authors would like to thank Dr. S. R. Ravishankar for his helpfuls comments on the manuscript.

[1] R. Marklein and P. Fellinger, ''Mathematisch-numerische Modellierung der Ausbreitung und Beugung akustischer Wellen,'' Proc. DGZfP-Seminar über Modelle und Theorien für Ultraschallprüfung, 5–6 November 1990, BAM Berlin.

[2] H. K. Versteeg and W. Malalaskera, *An Introduction to Computational Fluid Dynamics. The Finite Volume Method* (Longman Scientific & Technical, Essex, 1995).

[3] F. Ottitsch, ''Schwache mehrdimensionale Druckwellen in Rohrleitungen zufolge unstetiger Querschnittsänderungen am Beispiel der Eisenbahnaerodynamik,'' dissertation TU-Wien, Fakultät für Maschinenbau (1995).

[4] G. R. Harris, ''Review of transient field theory for a baffled planar piston,'' J. Acoust. Soc. Am. **70**, 10–20 (1981).

[5] L. Sulak, T. Bowen, A. E. Pifer, P. A. Polakos, H. Bradner, A. Parvalesca, A. V. Jones, and J. Learned, ''Experimental studies of the acoustic signature of proton beams traversing fluid media,'' Nucl. Instrum. Methods **161**, 203–217 (1979).

[6] A. Peiffer, B. Köhler, B. Hasch, and W. Enghardt, ''Sound radiation caused by heavy ions stopping in water and its possibilities for dose distribution monitoring in modern cancer therapy,'' Proc. World Congress on Ultrasonics, Berlin, 1995, pp. 1095–1098.

[7] S. Petzold, ''Optimierung einer transkanülar arbeitenden Dopplersonde, gerätetechnische Realisierung und Verifizierung ihrer Funktion,'' Diplomarbeit Technische Universität Dresden, Institut für Biomedizinische Technik (1996).

An experimental investigation of the propagation of sound over a curved, rough, rigid surface

James P. Chambers
National Center for Physical Acoustics, University of Mississippi, University, Mississippi 38677

Yves H. Berthelot
School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0405

(Received 16 May 1996; revised 29 January 1997; accepted 22 March 1997)

Small-scale laboratory experiments have been conducted to study the propagation of sound over a curved, rough, rigid, convex surface in the range $kR=275-550$, and $k\epsilon=0.09-0.18$, where k is the wave number, R the radius of curvature, and ϵ the characteristic roughness length scale of steep-sloped, densely packed scatterers. Experimental results are presented at 6, 8, and 12 kHz, with the receiver successively on the surface, along the line of sight behind the surface, and along a vertical axis in the shadow zone. At low frequencies ($f=6$ kHz, $k\epsilon=0.09$), a boundary wave caused by coherent multiple scattering develops near the surface, creeps into the deep shadow zone, and increases the sound-pressure levels measured in the deep shadow zone by up to 8 dB as compared to the levels measured in the absence of surface roughness. At high frequencies ($f=12$ kHz, $k\epsilon=0.18$), surface roughness causes incoherent scattering and decreases the measured sound-pressure levels by about 10 dB on the surface, in the shadow zone, and by 10 to 20 dB along the line of sight behind the apex. It is shown that incoherent scattering occurs mostly in the bright zone, before the apex, whereas coherent scattering (boundary wave growth) occurs all along the surface, even in the deep shadow. [S0001-4966(97)03707-7]

PACS numbers: 43.20.Fn, 43.28.Fp, 43.30.Hw, 43.50.Vt [LCS]

INTRODUCTION

The scattering of sound by rough surfaces with characteristic roughness length scales smaller than the acoustic wavelength has received considerable attention over the last four decades in the context of underwater acoustics and, to a lesser extent, atmospheric acoustics. The book by Ogilvy[1] gives a very comprehensive review of the state of the art in acoustic and electromagnetic scattering by rough surfaces. Similarly, wave diffraction by curved surfaces, with radius of curvature much greater than the wavelength, has been treated in great detail. (See, for instance, the work of Fock[2] for electromagnetic waves, Pierce[3] and Berry and Daigle[4] for acoustic waves.) There seems to be, however, neither data nor theory on the combined problem of low-frequency scattering and high-frequency diffraction of sound waves propagating over curved rough surfaces, even though such problems are frequently encountered in either atmospheric or ocean acoustics (e.g., shadowing behind a realistic berm, hill, or seamount). Another motivation for studying the sound field behind a curved rough surface is that the problem is analogous[3] to sound propagation over a flat, rough ground in an upwardly refractive atmosphere, a common occurence in outdoor sound propagation. Within this context, Medwin and Novarini[5] have given a somewhat qualitative description of what might be expected when refraction and bottom scattering are combined in underwater acoustic propagation, but to the authors' knowledge, no work has been published on the diffraction of sound by curved, rough surfaces. The objective of this paper is to report experimental results on the relative contributions of diffraction and scattering in the shadow zone, behind a curved, rough surface. This problem has potential relevance in the design of effective sound barriers and berms, and also in the propagation of sound over realistic topography in an upwardly refracting atmosphere.

In many practical cases, the ground surface is made of or covered by steep-sloped, densely packed scatterers upon which sound waves impinge at shallow grazing angles. Within this context, the seminal work of Twersky,[6] Tolstoy,[7,8] and, more recently, that of Lucas and Twersky[9,10] on sound propagation over a flat, rough surface is particularly important. The main result is that the coherent part of the scattered field appears in the form of a so-called boundary wave localized near the rough surface. The amplitude of this boundary wave can be very substantial at large distances in outdoor sound propagation near a rough ground because it spreads cylindrically and not spherically. Its amplitude is proportional to ϵk, where ϵ is the characteristic roughness length scale which will be defined in Sec. I, and where k is the acoustic wave number. (One should note that methods based on perturbation techniques cannot predict this first-order effect in ϵk because these methods are intrinsically second-order effects. For instance, the classical result[11] that incoherent scattering by surface irregularities causes the reflection coefficient to decrease by a factor of the order of $e^{-2k_z^2\epsilon^2}$ is a second-order effect in ϵk_z, where k_z represents the wave-number component normal to the surface.)

Scale model experiments with grazing sound propagation over flat, rigid, rough surfaces have been reported by Medwin and others.[12–14] They measured a boundary wave which propagates near the surface with cylindrical divergence at a slightly subsonic sound speed and, eventually, self-destructs at large distances because of incoherent scat-

tering. Medwin and D'Spain[14] have shown experimentally that, under proper conditions, surface roughness can cause an increase in the amplitude of the peak pressure by a factor of up to 4.5 (or 13 dB) as compared to what would have been measured over a smooth surface. Medwin *et al.*[13] also studied experimentally the diffraction of the boundary wave by a rigid 208° wedge covered with hemispherical scatterers. The main result is that the boundary wave diffracts in the same manner as a volume wave. At present, the authors are unaware of any experimental results (or theory) on the diffraction of sound by curved surfaces randomly covered with steep-sloped scatterers.

In this paper, we present data in the 6–12 kHz range on the propagation of sound over a convex surface (with radius of curvature $R=2.5$ m, i.e., $kR=275-550$), covered with small gravel with characteristic roughness length scale $\epsilon=0.085$ cm. With these parameters, the values of $k\epsilon$ range from 0.09 to 0.18, a range where substantial effects caused by surface roughness can be anticipated.[12–14] The experimental arrangement is described in the first section. In the second section, we present experimental data obtained successively on the curved rough surface, along the line of sight behind the surface, and along a vertical axis in the shadow zone. In an attempt to predict the measured data, we also present calculations of the creeping wave series[4,15] (for smooth convex surfaces) in which the surface impedance takes a finite value that is calculated according to the analogy between surface roughness and surface impedance. Such an analogy has been proposed by Tolstoy,[8,16] Howe,[17] and more recently by Attenborough and Taherzadeh.[18] Because of the limited success of this approach, more experiments have been carried out to further understand the relative contributions of scattering and diffraction of sound by curved, rough surfaces. These experiments are also reported in the second section. In the third section, we discuss how these indoor model experiments might scale for outdoor sound propagation applications and we draw some conclusions from the experimental work presented in this paper.

I. EXPERIMENTAL ARRANGEMENT

Experiments were conducted indoors in a standard laboratory environment. The curved surface was constructed of rigid plywood. Its radius of curvature was approximately 2.5 m near its apex. To produce rigid surface roughness, fishtank gravel was spread over a thin adhesive sheet of paper placed over the plywood surface. The statistics of the rough surface were measured and found to be comparable to those reported by Medwin[14] for #2 gravel. For a broad range of steep-sloped, isotropic, randomly rough, rigid surfaces, the characteristic roughness length scale ϵ is empirically described[14] in terms of the surface statistics: average height $\langle H\rangle$, standard deviation in height σ, and horizontal correlation length L_e, according to

$$\frac{\epsilon}{\langle H\rangle}=0.35-0.156\left[\sqrt{2}\left(\frac{\sigma}{L_e}\right)-1.5\right]^2. \tag{1}$$

We found that ϵ was 0.085 ± 0.025 cm in our experiment. The sound source was a small midrange tweeter driven by a tone-burst generator set to produce six cycles of a pure tone of tunable frequency. In our experiments, the frequency range of interest was between 6 and 12 kHz. The source was found to be omnidirectional to within 1 dB in a 15° cone around its axis of symmetry. The source was placed at a distance of 1.12 m from the apex of the convex surface. (See Fig. 1.) The receiving microphone was a 6.35-mm (1/4 in.) condenser microphone. The signal was amplified and filtered (3–30 kHz), captured on a digital oscilloscope triggered by the tone-burst generator, and transferred by general purpose interface bus (GPIB) to a computer for further analysis. The signal-to-noise ratio was 30 dB or better. In the results shown below, the measured quantity is the insertion loss (IL), which is, in decibels, the sound-pressure level in the presence of the rough surface relative to that which would have existed in a free field at the same location.

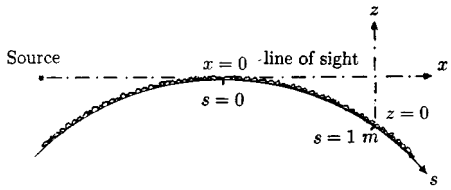


FIG. 1. Geometry of the problem.

II. RESULTS

A. Fully rough surface

The geometry of the problem is shown in Fig. 1. Experimental results were obtained at 6, 8, and 12 kHz, successively on the surface (Fig. 2), along the line of sight behind the surface (Fig. 3), and along a vertical axis in the shadow zone (Fig. 4). In Fig. 2, the abcissa is the arc distance s measured from the apex ($s=0$) such that s is positive in the shadow zone and negative in the bright zone. The solid line in Figs. 2–4 represents the theory for a curved, smooth, surface based on the creeping wave series.[4,15] (In this calculation, the receiver is 5 mm above the plywood surface to account for the physical size of the microphone used in the experiment.) As expected, there is a marked contrast between the effects of surface roughness at low frequencies and at high frequencies. Figure 2(a) shows that, at 6 kHz, the effect of surface roughness is to decrease the insertion loss (i.e., increase the sound-pressure level) relative to that of a smooth surface. For instance, at $s=100$ cm, the effect of surface roughness is to decrease the insertion loss by 6 dB relative to that measured on a smooth surface. At 12 kHz, shown in Fig. 2(c), incoherent scattering dominates and leads to increased insertion losses (i.e., reduced relative sound-pressure levels) both in the bright zone ($s<0$) and in the shadow zone ($s>0$). At 8 kHz, there is a transition regime where the insertion loss is higher in the bright zone ($s<0$) but lower than in the shadow zone past $s>60$ cm, relative to that which would exist over a curved smooth surface.

Figure 3 shows the insertion loss behind the curved surface along the line of sight, also at 6, 8, and 12 kHz. In Fig. 3(a)–(c), the abscissa x is zero at the apex of the curved surface. Clearly, the effect of surface roughness is to increase

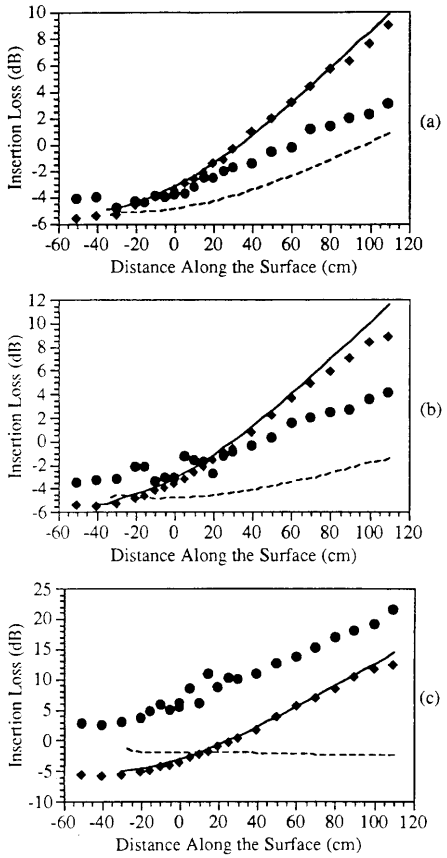


FIG. 2. Insertion loss on the surface at (a) 6 kHz, (b) 8 kHz, and (c) 12 kHz. ●: experimental data over the curved, rough surface; ◆: experimental data over the curved smooth surface; ———: creeping wave series for a rigid smooth surface; ----: creeping wave series for a smooth surface with impedance $Z^*=\rho c/(-ik\epsilon)$, with $\epsilon=0.085$ cm.

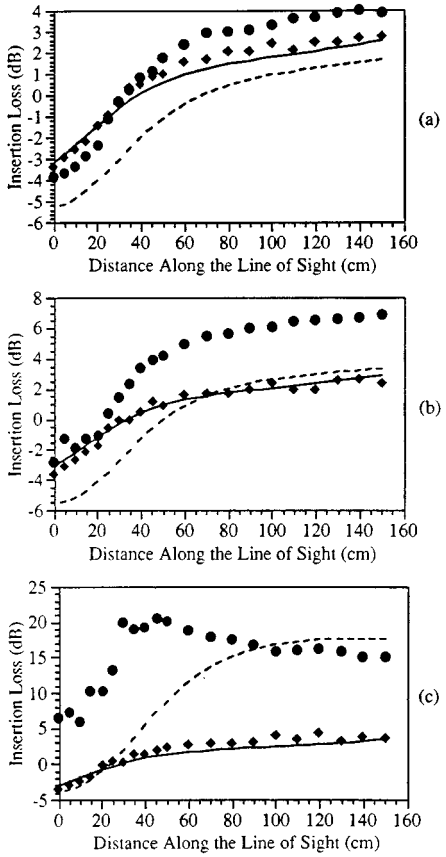


FIG. 3. Insertion loss along the line of sight at (a) 6 kHz, (b) 8 kHz, and (c) 12 kHz. ●: experimental data over the curved, rough surface; ◆: experimental data over the curved smooth surface; ———: creeping wave series for a rigid smooth surface; ----: creeping wave series for a smooth surface with impedance $Z^*=\rho c/(-ik\epsilon)$, with $\epsilon=0.085$ cm.

the insertion loss along the line of sight relative to that measured over the smooth surface. Figure 3(c) shows that, at 12 kHz, the effect of surface roughness is to increase the insertion loss measured along the line of sight, 40 cm behind the apex, by as much as 20 dB compared to what would have been observed over a smooth surface. In fact, the difference in relative insertion losses remains of the order of 15 dB or more from $x=40$ cm to $x=150$ cm, a rather astounding value considering the relatively small size of the roughness elements used in this experiment ($\epsilon k=0.18$, i.e., when $\epsilon \simeq 3\%$ of the wavelength).

Figure 4 shows the insertion loss behind a curved surface along a vertical axis in the shadow zone. The receiver's initial position was on the surface ($z=0$), deep into the shadow zone, at an arc distance $s=100$ cm from the apex. The receiver was moved vertically across the shadow zone, crossed the line of sight at $z=20$ cm, up to a height $z=40$ cm above the surface. At 6 kHz, Fig. 4(a) shows near the surface ($z<10$ cm) a pronounced decrease in insertion loss relative to that of a curved smooth surface. This decrease in insertion loss diminishes with increasing ϵk (or increasing frequency in our experiment) which is qualitatively consistent with the notion that a boundary wave creeps into the shadow zone.[13] Indeed, at 12 kHz the insertion loss of the curved, rough surface is higher than that of the smooth surface at all locations along the z axis. An interesting trend in the data can be observed. The insertion loss of the curved, rough surface is lower at the surface ($z=0$) than it is just above (at $z=2$ cm). This dip indicates that a boundary wave is present, even at 12 kHz where incoherent scattering is clearly the dominant effect.

Figures 2–4 indicate that the effect of coherent scattering (i.e., the boundary wave) is significant at 6 kHz, i.e., at $k\epsilon=0.09$. An experiment was carried out with very small surface roughness to assess the smallest value of $k\epsilon$ at which the boundary wave becomes noticeable. In this experiment, the plywood surface was covered with thin sheets of chiseled plexiglass used for fluorescent overhead fixtures, with a measured value of $\epsilon=0.008$ cm. At 5 kHz ($k\epsilon=0.0075$), the results followed within experimental errors the insertion loss obtained over a smooth, rigid surface. However, at 10 kHz ($k\epsilon=0.015$), the insertion loss for the rough surface was systematically slightly lower ($\sim$0.5 dB) than for the smooth rigid surface, thus showing that the effect of the boundary wave becomes measurable at values of ϵ/λ as low as 0.24%.

B. Analogy between surface roughness and surface impedance

Tolstoy[8,16] and Howe[17] have shown that, under certain conditions, one can rewrite the boundary condition above a flat, rough surface in the form of a standard impedance

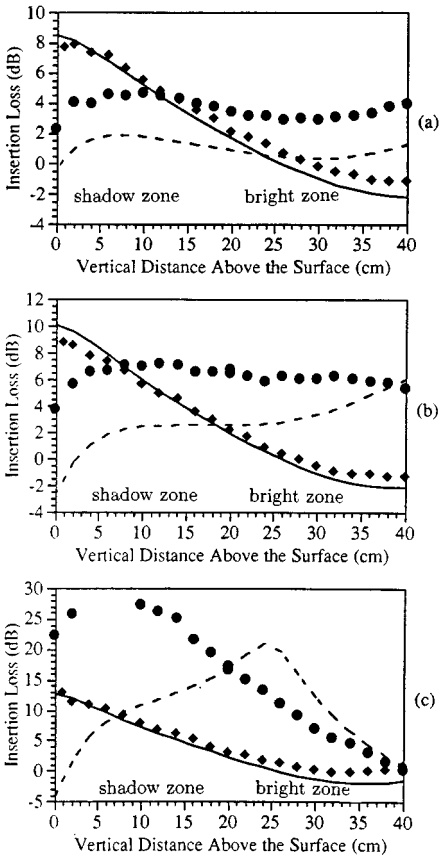


FIG. 4. Insertion loss along a vertical axis at an arc length $s=1.00$ m behind the apex on the surface. $z=0$ is on the surface, $z=20$ cm is the line of sight, $z=40$ cm is in the bright zone. (a) 6 kHz, (b) 8 kHz, and (c) 12 kHz. ●: experimental data over the curved, rough surface; ◆: experimental data over the curved smooth surface; ———: creeping wave series for a rigid smooth surface; ---: creeping wave series for a smooth surface with impedance $Z^*=\rho c/(-ik\epsilon)$, with $\epsilon=0.085$ cm.

boundary condition. The approach consists in replacing the rough boundary by a distribution of monopoles and dipoles whose scattering coefficients depend on the geometry of the boundary. If the scatterers are smaller than a wavelength, their strengths can be calculated from Rayleigh scattering theory. In the case of isotropic roughness, the boundary condition is written[8,16] as

$$\phi_z+\epsilon(k^2\phi+\sigma\phi_{zz})=0, \tag{2}$$

where ϕ is the acoustic velocity potential, the z axis is normal to the surface baseline (no curvature assumed), and σ is a parameter characterizing the dipole interactions. For the purpose of this paper, the scatterers are assumed to be perfectly rigid, although this is not a limitation of the approach. When $|\sigma\phi_{zz}|\ll|k^2\phi|$, i.e., when $\sigma\sin^2\theta\ll 1$, where θ is the grazing angle, the boundary condition (2) is approximated by

$$\phi_z+\epsilon k^2\phi=0. \tag{3}$$

This approximate boundary condition is therefore appropriate at near grazing sound propagation and/or for sparse scatterers. This boundary condition at a rough interface is in the form of a standard impedance boundary condition at a flat plane (with the $e^{-i\omega t}$ convention)

$$\phi_z+\eta\phi=0, \tag{4}$$

where $\eta=ik\beta$ and $\beta=\rho c/Z_s$ is the normalized surface admittance, ρc being the characteristic impedance of the upper fluid, and Z_s is the acoustic impedance of the smooth surface. By comparing boundary conditions (3) and (4), it appears that there is a formal analogy between sound propagation near a flat impedance plane and near a randomly rough, rigid plane, if $\eta=k^2\epsilon$, i.e., if

$$\beta=\frac{\rho c}{Z_s}=-ik\epsilon. \tag{5}$$

Therefore, rigid surface roughness can be understood as adding a purely reactive part to the surface impedance. This analogy has been suggested by several authors, and it has also been discussed in the context of nonrigid rough surfaces.[8,16,17] Recently, Attenborough and Taherzadeh[18] have combined and somewhat generalized Tolstoy and Howe's results. (In Attenborough and Taherzadeh's paper, ϵ appears as $\sigma_V/2$, where σ_V is the volume of scatterers per unit area.) With the analogy represented by Eq. (5), the boundary wave that propagates along a rough surface is formally equivalent to Wenzel's trapped surface wave[19] near an impedance plane with large reactance. Indeed, Wenzel's Eq. (20) in Ref. 19 has exactly the same form as Tolstoy's equation (13) in Ref. 20 if η is replaced by $k^2\epsilon$, and κ, the wave number of the boundary mode, by $\sqrt{k^2+\eta^2}$. Both boundary modes (surface wave or boundary wave) decay exponentially with height. The $1/e$ height is $(\epsilon k^2)^{-1}$, (or $1/\eta$); both modes are slightly subsonic, but the deviation from the free-field sound speed is a second-order effect and it can be safely ignored.

If such an analogy is also valid over a curved surface, one should be able to predict the insertion loss of a curved, rough surface by evaluating the creeping wave series[4,15] with an effective surface impedance $Z_s^*=\rho c/(-ik\epsilon)$. The nondimensional surface impedance parameter q which appears in the creeping wave series is therefore calculated as $q^*=(kl)(k\epsilon)$, where $l=(R/2k^2)^{1/3}$ is the characteristic diffraction length[2] associated with the curvature of the hill.

The results are plotted with a dashed line in Figs. 2–4. Clearly, the analogy yields poor predictions, except possibly at low frequencies, on the surface, in the shadow zone [Fig. 2(a)]. A likely explanation is that the analogy does not take into account incoherent scattering. Also, it may not hold for densely packed scatterers. In other words, the equivalent impedance Z_s^* should also include a resistive component associated with the attenuation of the boundary wave. Therefore, the characteristic roughness length scale ϵ should be treated as a complex quantity, even when dealing with rigid, rough surfaces. Preliminary results with a complex ϵ have been recently reported[21] but more work needs to be done to establish the range of validity of the analogy between a rough boundary and an impedance boundary with finite curvature.

C. Partially rough surface

Three additional experiments were carried out to determine the relative effect of boundary wave growth (coherent scattering) and attenuation (incoherent scattering) along and behind a curved, rough surface with a configuration identical

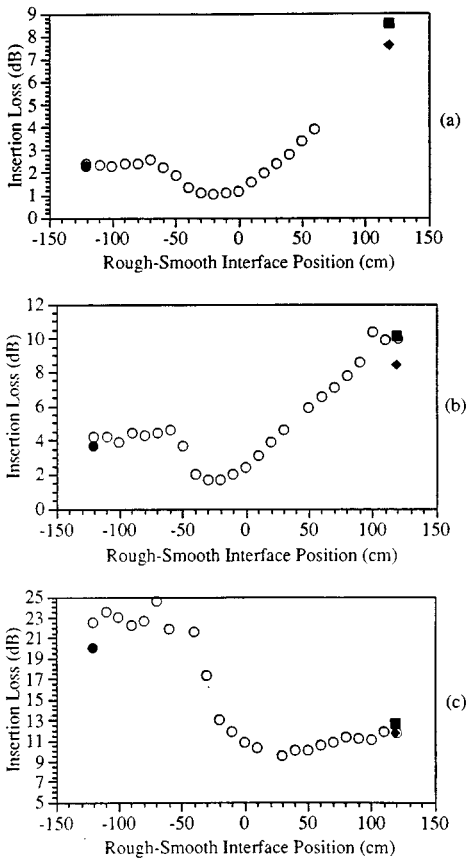


FIG. 5. Insertion loss in the deep shadow zone as roughness is gradually removed. The receiver is on the surface at an arc length, $s=1.00$ m behind the apex. The abcissa is the location of the smooth-rough interface. (a) 6 kHz, (b) 8 kHz, and (c) 12 kHz. ○: experimental data; ●: experimental data from Fig. 2 for the curved, fully rough surface; ◆: experimental data from Fig. 2 for the curved, fully smooth surface; ■: prediction from the creeping wave series for the fully smooth surface.

to that shown in Fig. 1. In the first two experiments, the source and receiver were fixed, the receiver being on the surface, deep into the shadow zone at $s=100$ cm. In the first experiment, strips of surface roughness (10-cm wide) were progressively removed from the bright side to the shadow side to study the transition from a fully rough to a fully smooth surface. In the second experiment, a single strip (20-cm wide) of roughness elements was progressively moved from the bright side to the shadow side, every 10 cm, over the otherwise smooth surface to examine where the scattering effects were dominant. In the third experiment, the bright side of the hill was left smooth, the shadow side was covered with roughness elements, and the insertion loss was measured successively on the surface and along the line of sight. The results are purely experimental.

1. Gradually removing the roughness (fixed source, fixed receiver)

Figure 5(a)–(c) shows the experimental results obtained at 6, 8, and 12 kHz, respectively, when gradually removing the roughness. The abscissa in Fig. 5 is the location of the smooth–rough interface, and the ordinate is the insertion loss measured at the receiver. (The data at 6 kHz is available only up to a smooth–rough interface located at $+60$ cm in the shadow zone, whereas it is plotted up to 120 cm for 8 and 12 kHz.) As roughness is removed, the insertion loss remains nearly constant until the smooth–rough interface reaches approximately $s=-60$ cm, s being the arc length measured from the apex. At this location, the ray path between the source and the receiver makes an angle of 8° with the horizontal axis (source apex), so that the angle of incidence with respect to the curved surface is about 22°. When roughness is removed beyond this point, the insertion loss in the shadow zone drops. The location of this drop characterizes where scattering effects dominate and that location appears to be nearly independent of frequency. However, the magnitude of this drop is much more dramatic at higher frequencies as one would expect, since attenuation by incoherent scattering is dominant at higher frequencies. As the roughness is removed near the apex and beyond, the insertion loss begins to rise. This is consistent with the notion that a boundary wave that is creeping into the shadow zone grows with distance. Removing roughness in the shadow zone reduces the distance over which the boundary wave can grow and one observes, at a given frequency, a progressive increase in insertion loss, as shown in Fig. 5.

2. Moving a strip of roughness across the smooth surface (fixed source, fixed receiver)

In this experiment, a strip of roughness with a width of 20 cm was moved over the otherwise smooth, curved surface, from the bright zone to the deep shadow zone. The source was fixed, 1.12 m from the apex. The receiver was in the deep shadow, on the surface, at an arc length $s=100$ cm. The results obtained at 6, 8, and 12 kHz are shown in Fig. 6(a)–(c), where the abcissa is the location of the center of the strip of roughness, and the ordinate is the insertion loss at the receiver. An increase in insertion loss can be observed at around $s=-60$ to -40 cm as the roughness strip is moved through the bright zone. This effect is more pronounced at higher frequencies and confirms that this region is where incoherent scattering effects are dominant. The drop in insertion loss (i.e., increase in sound-pressure level) that is observed for $s>-40$ cm confirms the notion that a boundary wave develops along the surface. Again, when comparing the insertion loss measured at different frequencies, one should keep in mind that the strip of roughness ''looks'' twice as long at 12 kHz than at 6 kHz, thus allowing for further boundary wave growth.

3. Roughness in the shadow zone only (fixed source, moving receiver)

In this experiment, the curved surface was completely smooth in the bright zone ($s<0$) and covered with gravel over the entire shadow zone ($s>0$). The source was 1.12 m from the apex. Figure 7(a)–(c) shows the insertion loss (open circles) measured on the surface at 6, 8, and 12 kHz, whereas Fig. 8(a)–(c) shows the insertion loss (open circles) along the line of sight, also at 6, 8, and 12 kHz. For comparison, we also show the insertion loss measured over the fully rough surface (black circles), over the fully smooth surface (black diamonds), as well as the predictions from the creeping wave series (solid line). Again, the results shown in Figs.

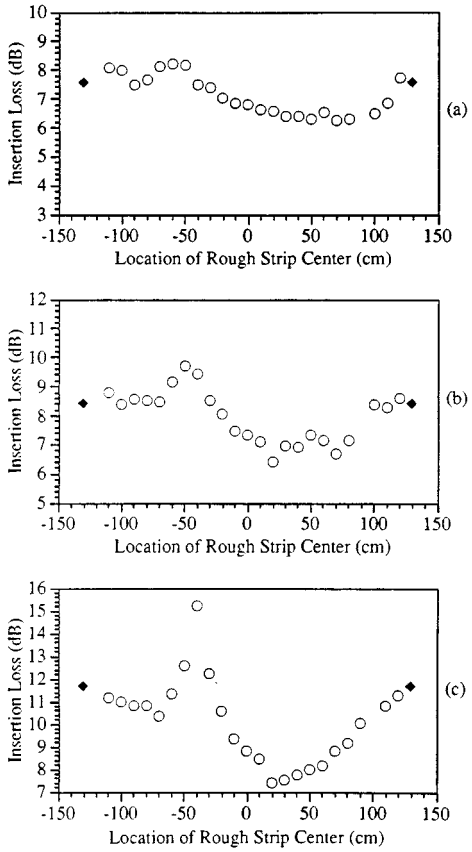


FIG. 6. Insertion loss in the deep shadow zone as a 20-cm-wide strip of roughness is gradually moved across the surface. The receiver is on the surface at an arc length, $s=1.00$ m behind the apex. The abcissa is the location of the center of the strip of roughness. (a) 6 kHz, (b) 8 kHz, and (c) 12 kHz. ○: experimental data; ◆: experimental data from Fig. 2 for the curved, fully smooth surface.

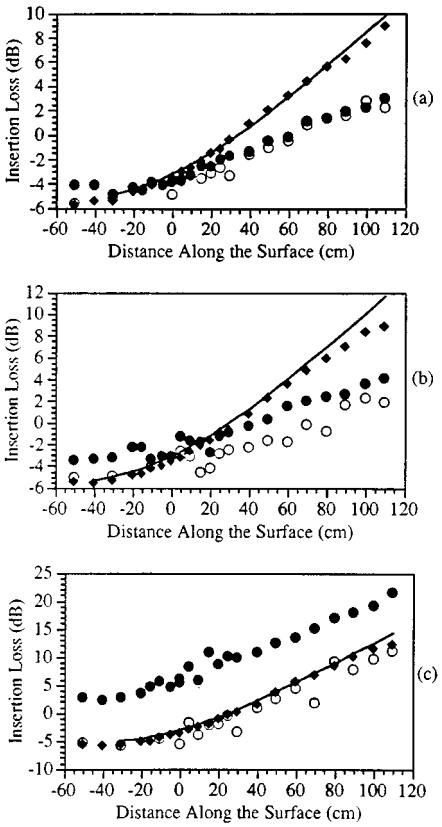


FIG. 7. Insertion loss on the curved surface at (a) 6 kHz, (b) 8 kHz, and (c) 12 kHz when the surface is fully smooth for $s<0$ and fully rough for $s>0$. The abcissa is the location of the receiver. ○: experimental data; ●: experimental data from Fig. 2 for the curved, fully rough surface; ◆: experimental data from Fig. 2 for the curved, fully smooth surface; ———: creeping wave series for a rigid smooth surface.

7 and 8 are consistent with the notion that boundary wave growth, which occurs near and along the surface in the shadow zone, is dominant at 6 kHz, whereas incoherent scattering, which occurs mostly in the bright zone, is the dominant effect at 12 kHz. Indeed, removing the scatterers in the bright zone consistently lowered the insertion loss at all frequencies and receiver locations (except at 6 kHz where scattering effects are negligible). This result may have some practical importance in the design of berms for highway noise reduction. To the authors' knowledge, no theory is currently available to predict the measured insertion losses.

III. SCALING

For practical applications, it is necessary to discuss the scaling involved in our indoor experiments. We discuss first the scaling associated with the diffraction problem, and then, the scaling associated with surface roughness. Fock[2] has shown that, for the diffraction of waves (wave number k) by a curved smooth surface of radius of curvature R, the characteristic nondimensional scale is

$$kl=\left(\frac{kR}{2}\right)^{1/3}. \tag{6}$$

The nondimensional field coordinates (ξ,ν) can be found[22] by introducing a parabolic coordinate system, (s,v), where s represents the arc length along the surface, and v is the height of the field point measured along a normal to the surface. The apex of the hill is defined by $s=0$ and $v=0$. The nondimensional parabolic coordinates which scale the problem are given by

$$\xi=\left(\frac{s}{R}\right)(kl), \quad \nu=\left(\frac{2v}{R}\right)(kl)^2. \tag{7}$$

Similarly, the distance x along the line of sight can be nondimensionalized into

$$X=2^{1/3}\left(\frac{x}{R}\right)(kl). \tag{8}$$

As far as the roughness scaling is concerned, one can introduce $H=kH$, $\overline{\sigma}=k\sigma$, and $L_e=kL_e$ as nondimensional variables for average height, standard deviation, or rms height, and correlation length, respectively, which leads[14] to the nondimensional roughness scale parameter $\overline{\epsilon}=k\epsilon$. For propagation over a flat, rough surface, the boundary wave is expressed in nondimensional form as

$$\left|\frac{P_B}{P_D}\right| \propto \overline{\epsilon}\, \overline{r}^{1/2} e^{-\overline{z}\,\overline{\epsilon}} e^{-\overline{\delta}\,\overline{r}} e^{-\overline{\beta}\,\overline{r}^2}, \tag{9}$$

where $\overline{r}=kr$, $\overline{z}=kz$ (z being the height above the boundary), $\overline{\delta}=\delta/k$, and $\overline{\beta}=\beta_M/k^2$, where δ and β_M are the linear and

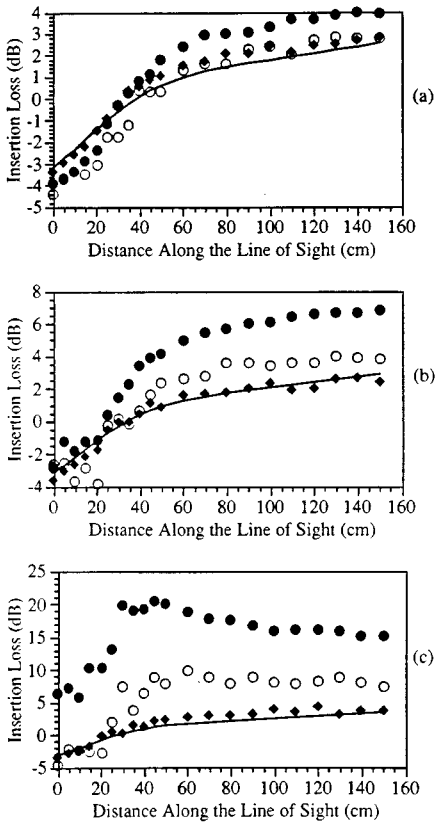


FIG. 8. Insertion loss along the line of sight behind the apex of the curved surface at (a) 6 kHz, (b) 8 kHz, and (c) 12 kHz when the surface is fully smooth for $s<0$ and fully rough for $s>0$. The abcissa is the location of the receiver along the line of sight. ○: experimental data; ●: experimental data from Fig. 3 for the curved, fully rough surface; ◆: experimental data from Fig. 3 for the curved, fully smooth surface; ———: creeping wave series for a rigid smooth surface.

quadratic attenuation coefficients introduced in Ref. 14. (β_M has the units of m^{-2}.) As an example, consider the case of a 100:1 scale factor. Figure 2(a) indicates that, by adding roughness of average height $\langle H\rangle \sim 25$ cm, rms height $\sigma \sim 16$ cm, and correlation length $L_e \sim 18$ cm (i.e., a roughness length scale parameter $\epsilon \sim 8.5$ cm), one should measure at 60 Hz a *relative increase of about 6 dB* in sound-pressure level in the shadow zone, on the surface at a distance of 100 m behind the apex of a hill of radius of curvature 250 m. Under the same conditions, it appears from Fig. 3(c) that, at 120 Hz, one should expect a relative decrease of about 15 dB in sound-pressure levels measured along the line of sight at a distance of 40 to 140 m behind the apex. These predictions are for rigid roughness over a rigid ground. Clearly, rigid roughness can drastically alter the sound field in the shadow zone and in the penumbra behind a convex obstacle, or, similarly, in an upwardly refractive atmosphere. Real scale outdoor experiments are needed to validate these results.

IV. CONCLUSIONS

Small-scale experiments have been conducted to study the propagation of sound over a curved, convex, rigid surface covered with steep-sloped, densely packed roughness elements. The sound-pressure levels (relative to free field) were measured at 6, 8, and 12 kHz, successively on the surface, along the line of sight behind the surface, and along a vertical axis in the shadow zone. It appears that surface roughness can have a drastic effect on the sound-pressure levels measured in the shadow zone and in the penumbra region. At low frequencies ($f=6$ kHz, $k\epsilon=0.09$), a boundary wave caused by coherent multiple scattering develops near the surface, creeps into the deep shadow zone, and increases the sound-pressure levels measured in the deep shadow zone by up to 8 dB as compared to the levels measured in the absence of surface roughness. At high frequencies ($f=12$ kHz, $k\epsilon=0.18$), surface roughness causes incoherent scattering and decreases the measured sound-pressure levels by about 10 dB on the surface, in the shadow zone, and by 10 to 20 dB along the line of sight behind the apex. Predictions obtained by modeling the rough surface as an equivalent smooth surface of finite impedance[17,18] in the computation of the creeping wave series[4,15] are, in general, poor estimates of the measured data, except on the surface and at low frequencies. (Recent results,[21] however, indicate that good predictions can be obtained if one introduces a complex ϵ whose phase accounts for the attenuation of the boundary wave by incoherent scattering.) To further understand the relative contributions of the diffraction and scattering mechanisms, additional experiments were performed with a partially rough, curved surface. It appears that incoherent scattering occurs mostly in the bright zone, before the apex, whereas coherent scattering (boundary wave growth) occurs all along the surface, even in the deep shadow. These effects may be important when considering long-range sound propagation over realistic terrain, or in the design of effective sound barriers and berms for highway noise reduction. At present, the authors are not aware of any theory available to predict the measurements reported in this paper.

ACKNOWLEDGMENTS

James P. Chambers was supported by a Hertz Fellowship from the Fannie and John Hertz Foundation for his graduate studies at the Georgia Institute of Technology. The authors are also very grateful to Shane Stone for his experimental assistance and to Lou Sutherland for his insightful comments and constructive criticism of this paper.

[1] J. A. Ogilvy, *Theory of Wave Scattering from Random Rough Surfaces* (Hilger, New York, 1991).
[2] V. A. Fock, *Electromagnetic Diffraction and Propagation Problems* (Pergamon, New York, 1965).
[3] A. D. Pierce, *Acoustics: An Introduction to Its Physical Principles and Applications* (McGraw-Hill, New York, 1981), pp. 469–478.
[4] A. Berry and G. Daigle, ''Controlled experiments on the diffraction of sound by a curved surface,'' J. Acoust. Soc. Am. **83**, 2047–2058 (1988).
[5] H. Medwin and J. C. Novarini, ''Modified sound refraction near a rough ocean bottom,'' J. Acoust. Soc. Am. **76**, 1791–1796 (1984).
[6] V. Twersky, ''On scattering and reflection of sound by rough surfaces,'' J. Acoust. Soc. Am. **29**, 209–225 (1957).
[7] I. Tolstoy, ''The scattering of spherical pulses by slightly rough surfaces,'' J. Acoust. Soc. Am. **66**, 1135–1144 (1979).
[8] I. Tolstoy, ''Smoothed boundary conditions, coherent low-frequency scatter, and boundary modes,'' J. Acoust. Soc. Am. **75**, 1–22 (1984).
[9] R. J. Lucas and V. Twersky, ''Coherent response to a point source irradiating a rough plane,'' J. Acoust. Soc. Am. **76**, 1847–1863 (1984).
[10] R. J. Lucas and V. Twersky, ''Inversion of data for near-grazing propagation over rough surfaces,'' J. Acoust. Soc. Am. **80**, 1459–1472 (1986).
[11] C. Eckart, ''The scattering of sound from the sea surface,'' J. Acoust. Soc. Am. **25**, 556–570 (1953).

[12] H. Medwin, J. Bailie, J. Bremhorst, B. J. Savage, and I. Tolstoy, ''The scattered acoustic boundary wave generated by grazing incidence at a slightly rough rigid surface,'' J. Acoust. Soc. Am. **66**, 1131–1134 (1979).
[13] H. Medwin, G. L. D'Spain, E. Childs, and S. J. Hollis, ''Low-frequency grazing propagation over periodic steep-sloped rigid roughness elements,'' J. Acoust. Soc. Am. **76**, 1774–1790 (1984).
[14] H. Medwin and G. L. D'Spain, ''Near-grazing, low-frequency propagation over randomly rough, rigid surfaces,'' J. Acoust. Soc. Am. **79**, 657–665 (1986).
[15] Y. H. Berthelot, ''A note on the acoustic penumbra behind a curved surface,'' J. Acoust. Soc. Am. **99**, 2428–2429 (1996).
[16] I. Tolstoy, ''Energy transmission into a shadow zone by rough surface boundary wave,'' J. Acoust. Soc. Am. **69**, 1290–1298 (1981).
[17] M. S. Howe, ''On the long range propagation of sound over irregular terrain,'' J. Sound Vib. **98**, 83–94 (1985).
[18] K. Attenborough and S. Taherzadeh, ''Propagation from a point source over a rough finite impedance boundary,'' J. Acoust. Soc. Am. **98**, 1717–1722 (1995).
[19] A. R. Wenzel, ''Propagation of waves along an impedance boundary,'' J. Acoust. Soc. Am. **55**, 956–963 (1974).
[20] I. Tolstoy, ''Rough surface boundary wave attenuation due to incoherent scatter,'' J. Acoust. Soc. Am. **77**, 482–491 (1985).
[21] Y. H. Berthelot and J. P. Chambers, ''On the analogy between sound propagation over a rough surface and sound propagation over a smooth surface with modified surface impedance,'' in *Proceedings of the 7th Long Range Sound Propagation Symposium* (Ecole Centrale de Lyon, Ecully, France, 1996).
[22] Y. H. Berthelot and J.-X. Zhou, ''Scale model experiments on the validity of the matched asymptotic expansions theory for sound diffraction by a curved surface of finite impedance,'' J. Acoust. Soc. Am. **93**, 605–608 (1993).

Focusing and imaging using eigenfunctions of the scattering operator

T. Douglas Mast[a)]
Department of Electrical Engineering, University of Rochester, Rochester, New York 14642

Adrian I. Nachman
Department of Mathematics, University of Rochester, Rochester, New York 14627

Robert C. Waag
Departments of Electrical Engineering and Radiology, University of Rochester, Rochester, New York 14627

(Received 15 October 1996; accepted 14 February 1997)

An inverse scattering method that uses eigenfunctions of the scattering operator is presented. This approach provides a unified framework that encompasses eigenfunction methods of focusing and quantitative image reconstruction in arbitrary media. Scattered acoustic fields are described using a compact, normal operator. The eigenfunctions of this operator are shown to correspond to the far-field patterns of source distributions that are directly proportional to the position-dependent contrast of a scattering object. Conversely, the eigenfunctions of the scattering operator specify incident-wave patterns that focus on these effective source distributions. These focusing properties are employed in a new inverse scattering method that represents unknown scattering media using products of numerically calculated fields of eigenfunctions. A regularized solution to the nonlinear inverse scattering problem is shown to result from combinations of these products, so that the products comprise a natural basis for efficient and accurate reconstructions of unknown inhomogeneities. The corresponding linearized problem is solved analytically, resulting in a simple formula for the low-pass-filtered scattering potential. The linear formula is analytically equivalent to known filtered-backpropagation formulas for Born inversion, and, at least in the case of small scattering objects, has advantages of computational simplicity and efficiency. A similarly efficient and simple formula is derived for the nonlinear problem in which the total acoustic pressure can be determined based on an estimate of the medium. Computational results illustrate focusing of eigenfunctions on discrete and distributed scattering media, quantitative imaging of inhomogeneous media using products of retransmitted eigenfunctions, inverse scattering in an inhomogeneous background medium, and reconstructions for data corrupted by noise. [S0001-4966(97)02308-4]

PACS numbers: 43.20.Fn, 43.60.Pt, 43.35.Wa, 43.80.Qf [ANN]

INTRODUCTION

This paper presents a new inverse scattering method that employs the focusing properties of certain acoustic fields obtained by retransmitting eigenfunctions of the scattering operator.

Eigensystem decomposition of the scattering operator, regardless of the inversion method employed, has potential advantages in methods of collecting and analyzing scattering data. Previous work in electrical impedance tomography has employed eigenfunction decomposition of an operator associated with the measurement process to determine optimal input current patterns and quantify the achievable resolution of imaging systems.[1,2] These optimal inputs can also be determined by iteratively retransmitting input patterns that are proportional to the measured scattered field. This approach is essentially an analog implementation of the ''power method'' for determining the eigenvectors of matrices.[2,3]

Likewise, the techniques of optical and acoustic phase conjugation[4–7] and the analogous process of time reversal[8,9] can be understood as analog methods of computing the eigenfunctions of an operator associated with the phase conjugation or time-reversal process. Simple focusing by phase conjugation, in which received echoes are conjugated or time reversed and retransmitted, is equivalent to a single iteration of the power method. Further iterations of this procedure correspond to additional steps in the power method, and thus converge to the most significant eigenfunction of the associated operator at a rate specified by the ratio of the two largest eigenvalues.[3] The eigenfunctions of the ''time-reversal operator,'' whether obtained by iterative time reversal or by numerical diagonalization, have been previously shown to correspond to source distributions that can focus incident energy on strong, pointlike scatterers.[10,11]

Eigensystem analysis has historically played a role in the theory of inverse scattering for radially symmetric objects.[12] For these objects, separation of variables naturally leads to a representation of the scattering operator in terms of trigonometric functions. Since these eigenfunctions are the same for any radial scatterer, the inverse scattering problem could be reduced to the problem of determining the unknown object from the eigenvalues of the scattering operator.

[a)]Current affiliation: Applied Research Laboratory, The Pennsylvania State University, P.O.B. 30, State College, PA 16802.

However, before the method presented here, the focusing properties of eigenfunctions have not been exploited for quantitative reconstruction of inhomogeneous media. It is stated in Ref. 9 that the concept of time reversal ''cannot be directly compared to computed tomography'' or to ''techniques that generate the image of the medium through signal analysis.'' Although the basic principles of focusing on point targets using the eigenfunctions of scattering operators have been put forth in Ref. 10, these principles have not previously been shown to apply to general distributed inhomogeneities. Furthermore, no general imaging method has hitherto been based on these principles.

The current method presents a solution to the imaging problem by bringing together recent results in the theory of focusing, diffraction tomography, and inverse problems to synthesize a unified framework for quantitative imaging of inhomogeneous media. Application of the method shows that focusing on distributed inhomogeneities can be achieved using eigenfunctions and also provides a technique for quantitative imaging of discrete and distributed inhomogeneities using focusing properties.

This method has several advantages over current inverse scattering methods. First, the eigenfunction formulation provides optimal bases for reconstruction of unknown media, so that inversions are performed with the minimum possible complexity. Second, the method is applicable to any scattering medium for which the total acoustic pressure associated with an incident plane wave can be estimated. Inverse scattering in inhomogeneous background media as well as iterative nonlinear inverse scattering can therefore be directly implemented. Third, part of the computation necessary for the inverse scattering algorithm can be performed by analog means using ideas from the power method.

The present approach also provides new understanding about existing methods of focusing and imaging. For simple scattering objects, the new method presented here reduces to a quantitative specification of focusing similar to that obtained by iterative phase conjugation or time reversal. The eigenfunctions of scattering operators are shown not only to focus on pointlike scatterers, as has been previously shown,[10,11] but also to concentrate incident energy in the vicinity of general, distributed inhomogeneities. The method also improves on previous approaches to focusing using eigenfunctions in that quantitative images of medium parameters are obtained simultaneously with optimal incident-wave distributions. For the case of weakly scattering objects, the method reduces to a simple inversion algorithm that is mathematically equivalent to the filtered backpropagation algorithm,[13–15] but is optimally tailored to the unknown scattering medium. The method reduces to a comparably simple and efficient formula for the case of weakly nonlinear inverse scattering.

Analysis given in Sec. I shows that eigenfunctions of scattering operators are equal to the acoustic fields of effective source distributions that are proportional to the compressibility contrast of the scattering object. An inverse scattering method that incorporates products of retransmitted fields of eigenfunctions is presented. The general method is then employed to derive an analytic inversion formula valid under the Born approximation as well as a simple nonlinear formula valid for small multiple-scattering effects. Numerical implementation of these methods is presented in Sec. II. Numerical results shown in Sec. III illustrate focusing on discrete and distributed inhomogeneities using a few eigenfunctions. Also, quantitative inverse scattering results are shown both within the context of a homogeneous background medium and an inhomogeneous background medium.

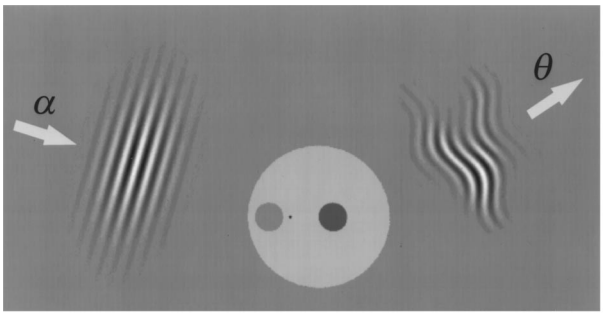


FIG. 1. Scattering configuration. An incident plane wave traveling in the direction $\boldsymbol{\alpha}$ is scattered by an inhomogeneity and the scattered field is measured in the direction $\boldsymbol{\theta}$.

I. THEORY

A. Background

An inverse scattering method for a medium of variable sound speed is derived. For simplicity of exposition, the derivation is given for the canonical two-dimensional scattering configuration sketched in Fig. 1. However, with minor modifications, the method is applicable to arbitrary geometries and dimensions.

When the incident pressure is a plane wave of unit amplitude propagating in the direction $\boldsymbol{\alpha}$, so that $p_i(\mathbf{x})=e^{ik\boldsymbol{\alpha}\cdot\mathbf{x}}$, the corresponding total acoustic pressure $p(\mathbf{x},\alpha)$ at the position $\mathbf{x}$ is given by the Lippman–Schwinger equation[16,17]

$$p(\mathbf{x},\alpha)=e^{ik\boldsymbol{\alpha}\cdot\mathbf{x}}-\int G_0(\mathbf{x}-\mathbf{y},k)q(\mathbf{y})p(\mathbf{y},\alpha)dy, \qquad (1)$$

where $G_0(\mathbf{x}-\mathbf{y})$ is the Green's function for the Helmholtz equation in a homogeneous medium. In an unbounded two-dimensional medium, $G_0(\mathbf{x}-\mathbf{y})$ is given by the Hankel function $(i/4)H_0^{(1)}(k|\mathbf{x}-\mathbf{y}|)$.[18] The angle α is defined as the angle corresponding to the direction unit vector $\boldsymbol{\alpha}$, the wave number k is equal to $2\pi f/c_0$ where c_0 is the wave speed of the background medium, and f is the temporal frequency of the incident wave. The integral appearing in Eq. (1), as well as subsequent integrals in x and y, are understood to be taken over the entire plane in $\mathbb{R}^2$. The scattering potential q is given for a medium of variable sound speed by

$$q(\mathbf{x})=-k^2\left(\frac{c_0^2}{c^2(\mathbf{x})}-1\right). \qquad (2)$$

The quantity within parentheses is equal, for a medium of constant density, to the compressibility contrast γ_κ, as defined in Ref. 17. The scattering potential is assumed to be

real-valued and to be short-range, that is, the potential q decreases at large distances such that

$$|q(\mathbf{x})| \leq C(1+|\mathbf{x}|)^{-1-\delta}, \tag{3}$$

where $|\mathbf{x}|$ is the magnitude of the position vector $\mathbf{x}$, for some $\delta > 0$.

At a measurement radius r in the far field and a measurement angle θ, the scattered pressure, $p_s = p - p_i$, is of the form

$$p_s(r,\theta,\alpha) = -\sqrt{\frac{i}{8\pi}}\,\frac{e^{ikr}}{\sqrt{kr}} A(\theta,\alpha) + o\left(\frac{1}{\sqrt{kr}}\right), \tag{4}$$

where A is the far-field pattern of the scattered pressure

$$A(\theta,\alpha) = \int e^{-ik\boldsymbol{\theta}\cdot\mathbf{x}} q(\mathbf{x}) p(\mathbf{x},\alpha)\, dx. \tag{5}$$

The incident pressure may be more generally taken as a superposition of plane waves propagating in all directions,

$$p_i(\mathbf{x}) = \int f(\alpha) e^{ik\boldsymbol{\alpha}\cdot\mathbf{x}}\, d\alpha. \tag{6}$$

The far-field pattern of the corresponding scattered acoustic pressure is then

$$Af(\theta) = \int A(\theta,\alpha) f(\alpha)\, d\alpha. \tag{7}$$

Equation (7) defines an operator A that maps an incident-wave distribution $f(\theta)$ into the corresponding far-field scattered pressure $Af(\theta)$. The operator A is related to the usual scattering operator S (Ref. 19) by

$$S = I - \frac{i}{4\pi} A, \tag{8}$$

where I is the identity operator.

The operator A is compact[19] and therefore has a countable number of discrete eigenvalues with zero as the only possible cluster point. In practice, only a finite number of eigenvalues are distinguishable from zero. Since the potential q is real-valued, the scattering operator is unitary, so that the eigenvalues of A lie in the complex plane on the circle centered at $-4\pi i$ and passing through the origin. It also follows that A is normal ($A^*A = AA^*$, where A^* is the Hermitian transpose of A), so that an orthonormal basis $\{f_i\}$ for $L^2[0,2\pi]$ exists consisting of eigenfunctions of A.

Since A is a normal operator, the Hermitian transpose A^* satisfies the relation $A^* f_i = \lambda^* f_i$, where f_i is an eigenfunction of A and λ_i^* is the complex conjugate of λ_i. The eigenfunctions of A therefore also satisfy the equation

$$A^*A f_i = |\lambda_i|^2 f_i. \tag{9}$$

Thus the functions f_i also constitute a basis of eigenfunctions for A^*A and the corresponding eigenvalues are the squared magnitudes of the eigenvalues of A. The operator A^*A is essentially a far-field analog of the ''time-reversal operator'' as defined in Ref. 10.

B. Focusing properties

The focusing properties of A are seen by considering the ratio of the scattered amplitude to the incident amplitude. Since A is normal, the magnitude of its largest eigenvalue is equal to the largest possible value of this ratio for any nonzero f:

$$|\lambda_1| = \sup\left[\frac{\|Af(\theta)\|_{L^2}}{\|f(\theta)\|_{L^2}}\right], \tag{10}$$

where $\sup(\cdot)$ denotes the least upper bound and $\|f(\cdot)\|_{L^2}$ denotes the root-mean-square magnitude of a square-integrable function. Thus the eigenfunction associated with the largest eigenvalue of A specifies an incident-wave distribution that maximizes the energy scattered to the far field. Other eigenfunctions also focus energy on inhomogeneities with an efficiency that is quantified by the associated eigenvalues.

The focusing property of eigenfunctions of A is further illustrated by introducing the acoustic fields of incident-wave distributions specified by the eigenfunctions. One may define retransmitted fields of an incident-wave distribution $f(\alpha)$ as

$$\begin{aligned} E(\mathbf{x}) &= \int f(\alpha) e^{ik\boldsymbol{\alpha}\cdot\mathbf{x}}\, d\alpha, \\ F(\mathbf{x}) &= \int f(\alpha) p(\mathbf{x},\alpha)\, d\alpha, \end{aligned} \tag{11}$$

where $E(\mathbf{x})$ is the retransmitted field associated with the incident-wave distribution in a homogeneous medium and $F(\mathbf{x})$ is the retransmitted field in a medium containing the inhomogeneity $q(\mathbf{x})$.

For incident-wave patterns corresponding to eigenfunctions that have nonzero eigenvalues, the retransmitted fields of Eqs. (11) can be written using Eqs. (5) and (7) in the form

$$\begin{aligned} E_i(\mathbf{x}) &= \frac{2\pi}{\lambda_i} \int J_0(k|\mathbf{x}-\mathbf{y}|) F_i(\mathbf{y}) q(\mathbf{y})\, dy, \\ F_i(\mathbf{x}) &= \frac{1}{\lambda_i} \int \langle p(\mathbf{x},\theta), e^{ik\boldsymbol{\theta}\cdot\mathbf{y}} \rangle F_i(\mathbf{y}) q(\mathbf{y})\, dy. \end{aligned} \tag{12}$$

The brackets in Eq. (12) denote the inner product

$$\langle u, v \rangle = \int_0^{2\pi} u(\theta) v^*(\theta)\, d\theta, \tag{13}$$

while the inner product appearing in the expression for E_i was evaluated using the identity

$$J_0(z) = \frac{1}{2\pi} \int e^{iz\cos\theta}\, d\theta \tag{14}$$

known as Parseval's integral.[20] The retransmitted fields of Eq. (11) are thus seen to be equivalent to a weighted convolution of the unknown scattering potential with inner products of acoustic fields.

When the scattering potential $q(\mathbf{x})$ is concentrated in a finite number of pointlike scatterers, each very small compared to a wavelength, Eq. (12) reduces to an expression of diffraction-limited focusing on each point scatterer. That is, for a scattering medium defined by

$$q(\mathbf{x})=\sum_{1}^{M}\mu_j\delta(\mathbf{x}-\mathbf{x_j}), \tag{15}$$

the retransmitted field $E_i(\mathbf{x})$ is

$$E_i(\mathbf{x})=\frac{2\pi}{\lambda_i}\sum_{j}F_i(\mathbf{x}_j)J_0(k|\mathbf{x}-\mathbf{x_j}|)\mu_j, \tag{16}$$

so that in this case, the retransmitted field $E_i(\mathbf{x})$ is equal to a weighted sum of Bessel functions, each centered at the location of one of the point scatterers. These Bessel functions correspond to a group of diffraction-limited main lobes, centered at each scatterer position $\mathbf{x_j}$, with corresponding Bessel sidelobes that combine coherently. Thus each retransmitted field E_i focuses to some extent on all of the individual point scatterers.

The close relationship between the retransmitted fields of eigenfunctions and the unknown scattering potential, as seen in Eq. (12), is an expression of the focusing property of eigenfunctions. That is, since eigenfunctions of A correspond to incident-wave patterns that concentrate energy within the support of the scattering potential, they can be said to focus on general distributed inhomogeneities as well as pointlike scatterers. This idea is illustrated numerically later in this paper.

C. Inverse scattering method

Because of the focusing properties outlined above, retransmitted fields of eigenfunctions are a useful starting point for inverse scattering reconstructions. A general inverse scattering method incorporating these ideas is outlined below.

The starting point for this method is an expression of the inverse scattering problem in terms of the operator A of Eq. (7) and the corresponding retransmitted fields defined in Eq. (11):

$$\langle Af_i,f_j\rangle=\delta_{ij}\lambda_i=\int F_i(\mathbf{x})E_j^*(\mathbf{x})q(\mathbf{x})dx, \quad i,j=1,2,\ldots . \tag{17}$$

The problem can be regularized by seeking the solution that minimizes the weighted L^2 norm

$$\|q\|_W^2=\int |q(\mathbf{x})|^2W(\mathbf{x})dx \tag{18}$$

with $W(\mathbf{x})$ an appropriate weight. For the analysis given below, this weight is defined as $W(\mathbf{x})=(1+|\mathbf{x}|)^{\delta},\delta>0$. For the explicit computations given later, other choices of $W(\mathbf{x})$ are more natural.

A solution to the minimization problem is obtained using the method of Lagrange multipliers, analogous to the approach used in Ref. 21 for a linearized electric impedance tomography problem. At a minimum, the (infinite-dimensional) gradient of $\|q\|_W^2$ is a linear combination of the gradients of the constraints in Eq. (17). The latter can be calculated using the two-potential formula[16]

$$A_{q_1}(\theta,\alpha)-A_{q_2}(\theta,\alpha)=\int p_1(\mathbf{x},\alpha)\times(q_1(\mathbf{x})-q_2(\mathbf{x}))p_2(\mathbf{x},\theta+\pi)dx, \tag{19}$$

where A_{q_1}, p_1, A_{q_2}, and p_2 are the scattering operators and the total acoustic pressures for the inhomogeneous media defined by $q_1(\mathbf{x})$ and $q_2(\mathbf{x})$, respectively. Equation (19) yields the derivative

$$\lim_{\epsilon\to 0}\frac{A_{q+\epsilon\tilde{q}}(\theta,\alpha)-A_q(\theta,\alpha)}{\epsilon}=\int p(\mathbf{x},\alpha)\tilde{q}(\mathbf{x})p(\mathbf{x},\theta+\pi)dx, \tag{20}$$

while the infinite-dimensional gradient of $\|q\|_W^2$ is found from

$$\lim_{\epsilon\to 0}\frac{\|q+\epsilon\tilde{q}\|_W^2-\|q\|_W^2}{\epsilon}=2\int q(\mathbf{x})\tilde{q}(\mathbf{x})W(\mathbf{x})dx. \tag{21}$$

The result follows that if the potential $q_M(\mathbf{x})$ solves the regularized inverse scattering problem [minimization of the weighted norm from Eq. (18) under the constraint of Eq. (17)], q_M must be of the form

$$q_M(\mathbf{x})=\frac{1}{W(\mathbf{x})}\sum_{l}\sum_{m}Q_{lm}F_l(\mathbf{x})\overline{F_m^*}(\mathbf{x}), \tag{22}$$

where $\overline{F_m^*}(\mathbf{x})$, the complex conjugate of the retransmitted field corresponding to an incoming condition at infinity, is defined as

$$\overline{F^*}(\mathbf{x})=\int f^*(\alpha)p(\mathbf{x},\alpha+\pi)d\alpha, \tag{23}$$

and the coefficients Q_{lm} are the Lagrange multipliers. If the above gradients are taken with respect to the real and imaginary parts of a complex potential, Eq. (22) as stated is also found to be valid when the potential q_M is complex. In some of the simplifying approximations made below, Eq. (22) will yield a complex potential q_M even when the data are assumed to come from a unitary scattering operator associated with the real potential q.

By substituting Eq. (22) into Eq. (17), the inverse problem is reduced to the problem of finding the coefficients Q_{lm} from the nonlinear system

$$\delta_{ij}\lambda_i=\sum_{l}\sum_{m}\left[\int\frac{F_i(\mathbf{x})E_j^*(\mathbf{x})F_l(\mathbf{x})\overline{F_m^*}(\mathbf{x})}{W(\mathbf{x})}dx\right]Q_{lm}, \quad i,j=1,2,\ldots , \tag{24}$$

where the dependence of the fields F and $\overline{F^*}$ on the scattering potential q is implicit.

In general, the scattering potential $q(\mathbf{x})$, and therefore the total pressure field $p(\mathbf{x},\alpha)$, are unknown in inverse scattering problems. The function $p(\mathbf{x},\alpha)$ that implicitly appears in Eq. (24) may therefore be replaced by the best available estimate for the total pressure. Equation (24) can then be solved for the coefficients Q_{lm} by standard numerical techniques for solution of linear systems.

The number of terms N can be chosen arbitrarily; however, increasing N beyond the number of nonzero eigenval-

ues of A is of limited benefit in reconstructions. For simple scattering objects, q can be represented by expansions employing small values of N. For instance, for an inhomogeneity consisting of finitely many point scatterers, N comparable to the number of scatterers is sufficient.

The above method simplifies further in the case of a weakly scattering medium, for which the total pressure p can be approximated by the incident pressure. In this case, taking the weight $W(\mathbf{x})\equiv 1$, the coefficients Q_{lm} can be evaluated analytically. From Eq. (22), under the Born approximation, the scattering potential takes the form

$$q_B(\mathbf{x})=\sum_l \sum_m Q_{lm}E_l(\mathbf{x})E_m^*(\mathbf{x}). \tag{25}$$

Substituting Eq. (25) into Eq. (5) gives the equation

$$A(\theta,\alpha)=\sum_l \sum_m Q_{lm}\int e^{-ik\boldsymbol{\theta}\cdot\mathbf{x}}E_l(\mathbf{x})E_m^*(\mathbf{x})p(x,\alpha)dx. \tag{26}$$

Replacement of $p(\mathbf{x},\alpha)$ in Eq. (26) by the incident plane wave $e^{ik\boldsymbol{\alpha}\cdot\mathbf{x}}$, use of Eq. (11), and integration in $\mathbf{x}$ over $\mathbb{R}^2$ yields

$$A(\theta,\alpha)=\frac{(2\pi)^2}{k^2}\sum_l \sum_m Q_{lm}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\delta(\boldsymbol{\theta}-\boldsymbol{\alpha}-\boldsymbol{\theta}'+\boldsymbol{\alpha}') \times f_l(\theta')f_m^*(\alpha')d\theta'\,d\alpha' \tag{27}$$

for $\theta-\alpha$ not equal to 0 or π. The double integral in Eq. (27) can be evaluated using the change of variables

$$x_1'=\cos\,\theta'-\cos\,\alpha', \quad x_2'=\sin\,\theta'-\sin\,\alpha', \tag{28}$$

which is one-to-one when restricted to the regions $\alpha'<\theta'$ and $\alpha'>\theta'$. Evaluation of the integral yields

$$|\sin(\theta-\alpha)|A(\theta,\alpha)=\frac{(2\pi)^2}{k^2}\sum_l \sum_m Q_{lm}(f_l(\theta)f_m^*(\alpha) + f_l(\alpha+\pi)f_m^*(\theta+\pi)). \tag{29}$$

Equation (29) can be solved for the coefficients Q_{lm} using the fact that the eigenfunctions $f_l(\theta)$ are orthonormal as well as the reciprocity identity[16]

$$A(\theta+\pi,\alpha+\pi)=A(\alpha,\theta). \tag{30}$$

The solution is

$$Q_{lm}=\frac{k^2}{8\pi^2}\int\int|\sin(\theta-\alpha)|A(\theta,\alpha)f_l^*(\theta)f_m(\alpha)d\alpha\,d\theta. \tag{31}$$

Equations (25) and (31) specify a solution q_B to the linearized inverse problem. This solution is, in general, complex, even when the true potential q is purely real. A physical way to understand why the Born approximation yields a complex scattering potential for a lossless medium is to recognize that this approximation neglects multiple scattering and thus, the resulting output energy differs from the input energy. The corresponding scattering operator is then no longer unitary, and is only physically realizable by a potential with a nonzero imaginary part. For weak scattering, the energy discrepancy is small and so is the imaginary part of the potential.

The analytic solution of Eqs. (25) and (31) is equivalent to the well-known filtered backpropagation formula[13–15] and has the advantage of computational simplicity, as discussed later in this paper. Equivalence between the two formulas is shown by formulating an expansion of $e^{-ik\boldsymbol{\alpha}\cdot\mathbf{x}}$, viewed as a function of α, in terms of the orthonormal basis $\{f_m(\alpha)\}$. In view of Eqs. (11), this expansion yields the identity

$$e^{-ik\boldsymbol{\alpha}\cdot\mathbf{x}}=\sum_m E_m^*(\mathbf{x})f_m(\alpha). \tag{32}$$

Substituting Eq. (31) in Eq. (25) and using Eq. (32) as well as its conjugate gives

$$q_B(\mathbf{x})=\frac{k^2}{8\pi^2}\int\int|\sin(\alpha-\theta)|A(\theta,\alpha)e^{ik\mathbf{x}\cdot(\boldsymbol{\theta}-\boldsymbol{\alpha})}d\alpha\,d\theta, \tag{33}$$

which is the standard filtered backpropagation formula. Equation (33) yields the low-pass-filtered version of the true potential q if multiple scattering effects are negligible. The correct nonlinear generalization of the linearized low-pass filtered solution q_B is the minimal L^2 (or weighted L^2) solution q_M, which is of a form specified by Eq. (22).

The inverse scattering method developed above can also be used with any orthonormal set of basis functions for $L^2[0,2\pi]$. For instance, reconstructions can be performed using eigenfunctions of A for axisymmetric objects rather than using the eigenfunctions associated with the measured A. In this case, the eigenfunctions take the form

$$f_m(\theta)=\frac{1}{\sqrt{2\pi}}e^{im\theta}, \quad m=0,\pm 1,\pm 2,\ldots\,. \tag{34}$$

The retransmitted fields E_m can be analytically evaluated to be

$$E_m(r,\phi)=\sqrt{2\pi}i^m e^{im\phi}J_m(kr), \tag{35}$$

and the coefficients Q_{lm} for the low-pass-filtered reconstruction of q are given by

$$Q_{lm}=\frac{k^2}{16\pi^3}\int\int|\sin(\theta-\alpha)|A(\theta,\alpha)e^{-il\theta}e^{im\alpha}d\alpha\,d\theta. \tag{36}$$

While the retransmitted fields specified by Eq. (35) are not ideally matched to nonaxisymmetric scattering media, they can be analytically evaluated and stored for use in fast reconstructions. Since these retransmitted fields are also unaffected by uncertainties in scattering measurements, they are suitable for reconstructions from data corrupted by noise.

Finally, use of the eigenfunction method beyond linear inversion is demonstrated by considering the case where the inhomogeneous-medium retransmitted fields F can be estimated from a first approximation to the scattering potential q. One approach in this case is to solve the full system of equations defined by Eq. (24); however, a more numerically efficient correction to the Born approximation can be obtained by invoking the localized nonlinear approximation introduced in Ref. 22 for electromagnetic scattering. This approximation follows from writing the Lippman–Schwinger equation [Eq. (1)] in the form

$$p(\mathbf{x},\alpha)=\Gamma(\mathbf{x})\left(e^{ik\boldsymbol{\alpha}\cdot\mathbf{x}}-\int (p(\mathbf{y},\alpha)-p(\mathbf{x},\alpha))\right.$$

$$\left.\times q(\mathbf{y})G_0(\mathbf{x}-\mathbf{y})dy\right), \tag{37}$$

where the quantity $\Gamma(\mathbf{x})$, called the depolarization tensor in electromagnetic scattering,[22] is defined by

$$\Gamma(\mathbf{x})=\left(1+\int q(\mathbf{y})G_0(\mathbf{x}-\mathbf{y})dy\right)^{-1}. \tag{38}$$

The second term in Eq. (37) is presumed to be small because the singularity of the Green's function is cancelled by the difference term appearing in the integrand. Thus the total pressure may be approximated by the formula

$$p(\mathbf{x},\alpha)\approx\Gamma(\mathbf{x})e^{ik\boldsymbol{\alpha}\cdot\mathbf{x}}. \tag{39}$$

The form for the scattering potential given by Eq. (22) then becomes

$$q_M(\mathbf{x})\approx\frac{\Gamma(\mathbf{x})^2}{W(\mathbf{x})}\sum_l\sum_m Q_{lm}E_l(\mathbf{x})E_m^*(\mathbf{x}). \tag{40}$$

Substituting this form into Eq. (5) and using Eq. (39) gives

$$A(\theta,\alpha)\approx\sum_l\sum_m Q_{lm}\int e^{-ik\boldsymbol{\theta}\cdot\mathbf{x}}W(\mathbf{x})^{-1}E_l(\mathbf{x})E_m^*(\mathbf{x})$$

$$\times\Gamma(\mathbf{x})^3e^{ik\boldsymbol{\alpha}\cdot\mathbf{x}}\,dx. \tag{41}$$

An approximate nonlinear formula for the scattering potential q can be obtained by taking $W(\mathbf{x})\equiv\Gamma(\mathbf{x})^3$. Equation (41) then yields the coefficients Q_{lm} from Eq. (31) and the resulting solution for the scattering potential is

$$q_M(\mathbf{x})\approx\sum_l\sum_m Q_{lm}\frac{E_l(\mathbf{x})E_m^*(\mathbf{x})}{\Gamma(\mathbf{x})}. \tag{42}$$

The solution of Eq. (42) is simplified by making the further approximation

$$\frac{1}{\Gamma(\mathbf{x})}\approx 2-\Gamma(\mathbf{x}), \tag{43}$$

which is valid for small scattering potentials. This substitution results in

$$q_M(\mathbf{x})\approx\sum_l\sum_m Q_{lm}(2E_l(\mathbf{x})-F_l(\mathbf{x}))E_m^*(\mathbf{x}). \tag{44}$$

This nonlinear equation for the potential q_M can be approximately solved by using a form of the retransmitted field $F_l(\mathbf{x})$ corresponding to the low-pass-filtered potential q_B or to another estimate of the scattering potential.

II. COMPUTATIONAL METHODS

The focusing and imaging methods outlined in Sec. I were implemented using numerically computed scattered fields of inhomogeneous objects. Scattering operators were calculated using a method due to Kirsch and Monk,[23] in which an inner solution of the Helmholtz equation for a medium of variable sound speed is matched to an outer solution of integral equations that implicitly satisfy the Sommerfeld radiation condition. The inner solution is obtained using a finite-element method, while the outer integral equations are solved using Nyström's method.[23]

Scattering data were calculated numerically for a number of incident plane waves evenly distributed over M angles between 0 and 2π. For each incident-wave angle, the scattered field was computed at M far-field receiver angles between 0 and 2π, so that the angular sampling rate was $M/(2\pi)$ samples per radian. The number of receiver angles M should be chosen such that the scattered field has no significant frequency components above the Nyquist frequency of $M/(4\pi)$ samples per radian. This computation yields a discrete representation of the scattering operator A as an $M\times M$ matrix, A_M.

The eigenfunctions of A and their associated eigenvalues were estimated numerically by direct computation of the eigenvectors and eigenvalues of A_M. Retransmitted fields of eigenfunctions were evaluated numerically by numerical integration of Eq. (11). Images of inhomogeneous objects were then obtained using a straightforward numerical implementation of Eqs. (25) and (31). The integrals appearing in these equations were evaluated using corresponding discrete summations of the components of A_M and its eigenvectors. For comparison, standard diffraction tomography inversions were also performed by numerical integration of Eq. (33).

Stability of the eigenfunction imaging method was tested by inversion of noisy data obtained by adding numerically generated Gaussian white noise to the scattering matrix A_M. The rms amplitude of the noise was specified as a fraction of the rms value of A_M. Thus, for instance, a signal-to-noise ratio of 6 dB was obtained by adding noise with an rms amplitude one-half the rms value of A_M.

Inversions were also performed using the basis of eigenfunctions corresponding to axisymmetric inhomogeneities. In this case, the formula of Eq. (25) was implemented numerically using the trigonometric basis functions defined in Eq. (34), the retransmitted fields given in closed form in Eq. (35), and the coefficients defined in Eq. (36).

Nonlinear eigenfunction images were obtained using the analytic formula of Eq. (44) with the total pressure $p(\mathbf{x},\alpha)$ approximated by the total pressure for a medium including a cylinder of specified radius and compressibility contrast. This computation employed an exact solution for the scattering of a plane wave by a cylinder.[17]

III. NUMERICAL RESULTS

Focusing of eigenfunctions on a distributed scattering object is illustrated in Fig. 2. Here, the magnitudes of the retransmitted fields $E_1(\mathbf{x})$ and $E_2(\mathbf{x})$ are shown for an inhomogeneity consisting of a weakly scattering triangle ($\gamma_\kappa=0.01$) approximately two wavelengths in height. The triangle is shown in outline together with the retransmitted fields. The corresponding scattering operator, calculated using the finite-element/Nyström method described above, was represented by a matrix of size 128×128. The retransmitted fields show that the significant eigenfunctions of A specify incident-wave patterns that concentrate energy within the support of the inhomogeneity. Notable is that this focused energy is distributed throughout the support of the triangle.

Implementation of the eigenfunction method in focusing on pointlike scatterers is illustrated in Figs. 3 and 4. These figures, obtained using the linearized eigenfunction method, show not only diffraction-limited focusing but also quantitative reconstructions of the associated scattering potentials. Figure 3 shows images made from the scattered field of two pointlike scatterers at locations $(-0.5,0)$ and $(0,-0.2)$, each of radius 0.01 and compressibility variation -0.9. The numerically computed scattered field was sampled at 128 equally spaced angles for each of 128 incident-wave angles, so that the operator A was represented by a 128×128 matrix. The wave number used was 10, so that the scatterers were separated by approximately one wavelength. Since, in this case, two eigenvalues of A were much larger than the remaining eigenvalues, the basic reconstruction required only the use of two retransmitted fields. This result illustrates that, for an inhomogeneity consisting of finitely many pointlike scatterers, the present inverse scattering method provides an accurate reconstruction with diffraction-limited point resolution using a corresponding number of eigenfunctions.

A stability test of the eigenfunction method is illustrated in Fig. 4. This image shows a reconstruction of the two pointlike scatterers of Fig. 3 using the same scattering data with added Gaussian white noise for a signal-to-noise ratio of 3 dB. The method of reconstruction was identical to that used for Fig. 3. The reconstruction shown is almost indistinguishable from the noiseless reconstruction, indicating the stability of the eigenfunction imaging method.

Linear eigenfunction images of the triangular inhomogeneity of Fig. 2 are presented in Fig. 5. These images were constructed using the same scattering data as that used for Fig. 2. The first image, obtained using five retransmitted fields, shows that strong focusing is achieved using only a few eigenfunctions of A. The entire inhomogeneity is wellinsonified and little incident energy is transmitted outside the support of the inhomogeneity. The second image, obtained using 15 eigenfunctions, shows that the eigenfunction method rapidly converges to the ideal low-pass-filtered solution for the scattering potential. Notable is that the eigenfunction method using 15 eigenfunctions required 69.1 s of CPU time on a Sun SPARCstation 10, while an analogous image obtained using the diffraction tomography formula of Eq. (33), with the integrals evaluated in an analogous manner, required 3014.3 s.

Eigenfunction reconstructions of a test phantom, shown in Figs. 6–8, illustrate application of the eigenfunction imaging method to a larger-scale imaging problem. The phantom, also represented in Fig. 1, is a cylinder of compressibility contrast 0.01 and diameter of 5 mm. Internal objects include a water-filled (cystic) region of diameter 1 mm, a wire of diameter 0.1 mm and compressibility contrast -0.5, and an internal cylinder of diameter 1 mm and compressibility contrast -0.01. Scattered fields were calculated using the methods described above, with the operator A discretized as a matrix of 256×256 points. The first image shown in Fig. 6, obtained using the single wave number $k=10$ has high resolution but contains ringing (Gibbs phenomenon) artifacts and loss of contrast in the cystic region. These artifacts are removed by compounding of images obtained using five linearly spaced wave numbers, $8\leqslant k\leqslant12$, so that the dimensionless parameter ka varied between 20 and 30. The five-frequency image, shown in the second panel of Fig. 6, also shows increased point and contrast resolution compared to the single-frequency image. Both images shown in Fig. 6 were obtained using the linearized eigenfunction method described above, with 64 eigenfunctions of A for $k=8$, 9, and 10, 68 eigenfunctions for $k=11$, and 72 eigenfunctions for $k=12$.

Reconstructions of the test phantom obtained from noisy data are shown in Fig. 7. Gaussian white noise was added to the scattering data employed in Fig. 6, so that the signal-to-noise ratio was 6 dB at each of the frequencies employed. The reconstructions employed the formula of Eq. (25) and coefficients obtained from Eqs. (34)–(36). The numbers of basis functions employed were equal to the number of eigenfunctions employed in Fig. 6. These results indicate the stability of the method for large objects with scattering data severely degraded by noise.

Nonlinear reconstructions of the same test phantom, obtained using Eq. (44), are presented in Fig. 8 together with linear reconstructions. In the nonlinear reconstructions, the retransmitted fields $F_l(\mathbf{x})$ were estimated using pressure fields associated with a cylinder of diameter 5 mm and compressibility contrast 0.01.[17] The scattering data employed was identical to that used in Fig. 6(b), with five linearly spaced wave numbers such that $20\leqslant ka\leqslant30$. The number of eigenfunctions employed in each image were also the same as those used for the images in Fig. 6. The first panel shows the real part of the nonlinear reconstruction, taken along the line $y=0$, together with the real part of the analogous linear reconstruction from Fig. 6(b). The nonlinear reconstruction shows improved resolution over the linear reconstruction by increased height of the peak associated with the internal wire. The second panel shows the imaginary part of the nonlinear reconstruction with the corresponding linear reconstruction from Fig. 6(b). Here, the inaccuracy of the Born approximation results in a significant imaginary part for the linear reconstruction, while the true potential is purely real. The nonlinear inversion shows improved quantitative accuracy over the linear inversion by reduction of the reconstructed imaginary part.

IV. DISCUSSION

Our method has shown that eigenfunctions of the scattering operator can be employed to focus on distributed inhomogeneities as well as pointlike scatterers. However, the focusing on distributed inhomogeneities occurs in a different manner from focusing on pointlike scatterers. That is, the incident energy is not maximized at a single point within the medium. Instead, when combined according to the present reconstruction method, retransmitted eigenfunctions specify incident-wave distributions that distribute energy throughout the inhomogeneous region. This type of focusing, which results from the eigenfunction property of maximizing the scattered energy, is clearly connected to imaging of the medium by inverse scattering.

The quantitative inverse scattering method presented in this paper can considerably simplify imaging computations.

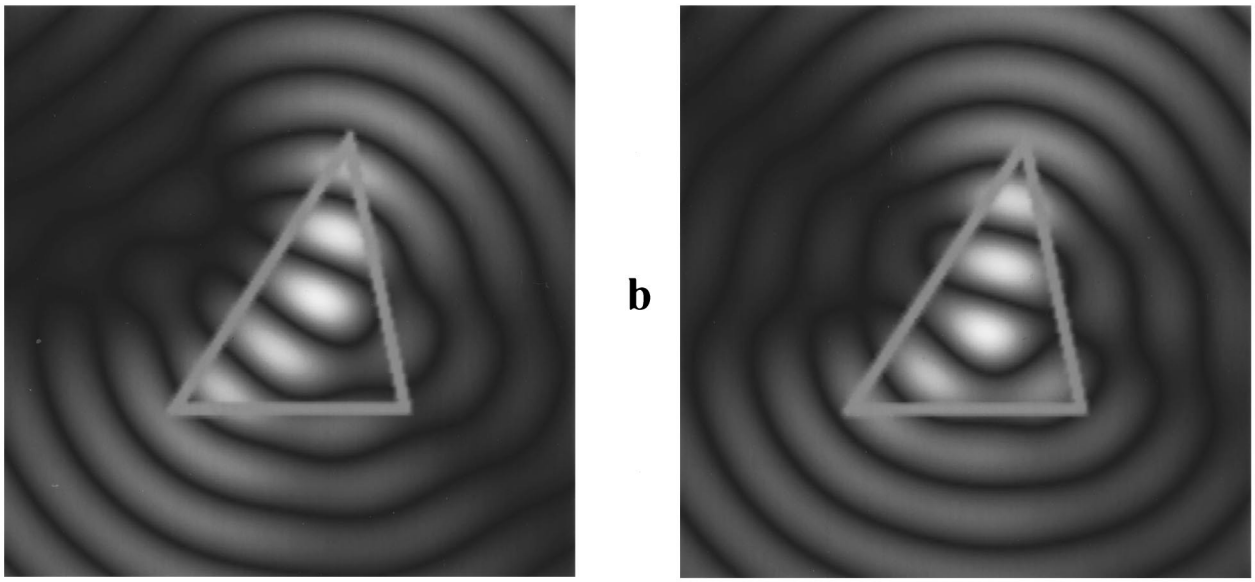


FIG. 2. Focusing on a distributed inhomogeneity. Magnitudes of the retransmitted fields of the two most significant eigenfunctions are shown on a linear gray scale with black indicating zero and white indicating maximum amplitude. The scattering object is a uniform triangle, compressibility contrast 0.01, within the sketched boundaries. (a) $E_1(\mathbf{x})$, (b) $E_2(\mathbf{x})$.

The computational complexity of the current method depends mainly on the number of significant eigenfunctions, which in turn depends only on the complexity of the scattering medium. Furthermore, the basis for expansion of the unknown medium is determined directly from the scattering data. Since the basis functions are closely related to the unknown medium, reconstructions performed using this basis employ a minimal amount of unnecessary information. This property gives the present inverse scattering method advantages over other methods in which a fixed basis is used to expand the unknown medium.[24–27] These advantages are most apparent for inhomogeneities a small number of wavelengths in size.

The present inverse scattering method also has the advantage of applicability to any medium for which the background pressure field can be estimated. Use of background pressure estimates can greatly improve accuracy over reconstructions based on simpler approximations. For instance, Born inversion can yield a spurious reconstructed imaginary part even when the true potential is real-valued; use of an estimated background pressure field can greatly reduce this error, as seen in Fig. 8. The inverse scattering method pre-

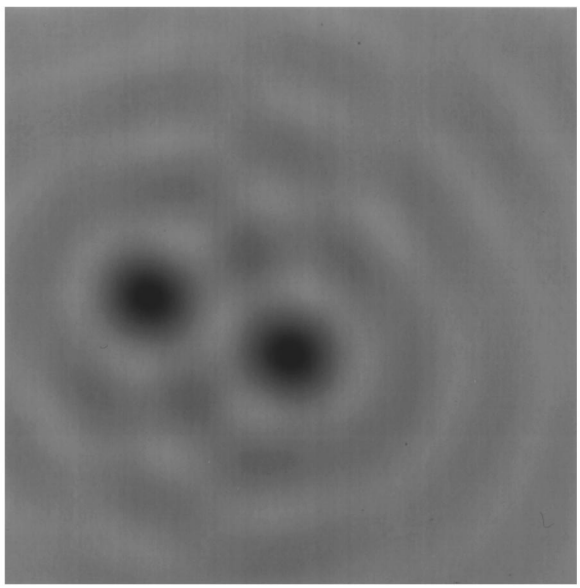


FIG. 3. Eigenfunction image of two pointlike scatterers, compressibility contrast -0.9, separated by approximately one wavelength. The image was obtained using retransmitted fields of the two most significant eigenfunctions.

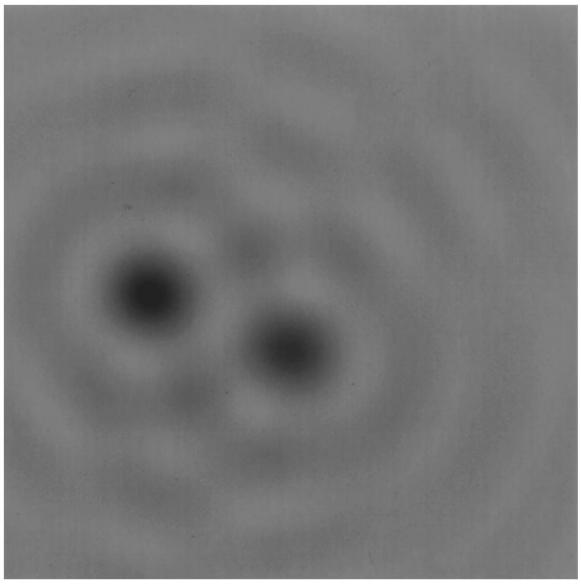


FIG. 4. Effect of noise on eigenfunction reconstruction. The object of Fig. 3 was reconstructed from two eigenfunctions of synthetically noised scattering data with a signal-to-noise ratio of 3 dB.

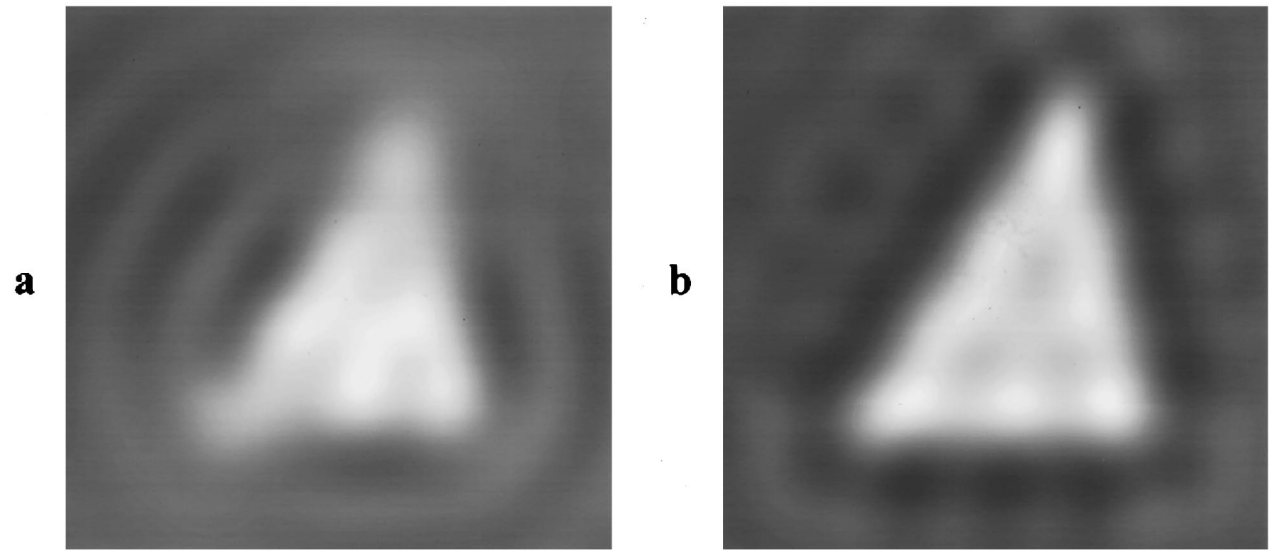


FIG. 5. Eigenfunction images of a triangle about three wavelengths in height having compressibility contrast 0.01. (a) Real part of an inversion obtained using retransmitted fields of five eigenfunctions. (b) Analogous image obtained using 15 eigenfunctions.

sented here is also extensible to any background medium for which a pressure field can be estimated, including moving fluids, layered or stratified media, and enclosed or otherwise bounded regions.

The imaging, focusing, and inverse scattering methods presented here also intrinsically take advantage of any potential increase in resolution that is associated with multiple scattering or other higher-order effects, as long as these effects are taken into account in the estimated pressure field. This increase in resolution is possible because the retransmitted fields of eigenfunctions may have desirable qualities, such as higher spatial-frequency components, that are associated with the presence of an inhomogeneous background. Such improvements in resolution have been shown experimentally for time-reversal focusing in a multiply scattering medium.[28]

The most significant eigenfunctions of A can be determined experimentally through iterative retransmission of received scattered fields in a manner similar to that performed in Refs. 1 and 10. This implementation of the power method[3] allows computation of the most significant eigenfunctions of A by analog means, which may be preferable to numerical computation for very large scattering objects. These eigenfunctions are useful as optimal incident-wave patterns for inverse scattering experiments.

The inverse scattering method presented here includes the assumption that the scattering potential is purely real, that is, the inhomogeneous medium is assumed to have zero absorption. Eigensystem analysis of the scattering operator A is more complicated in the presence of absorption.[29,30] However, the methods introduced here are expected to still

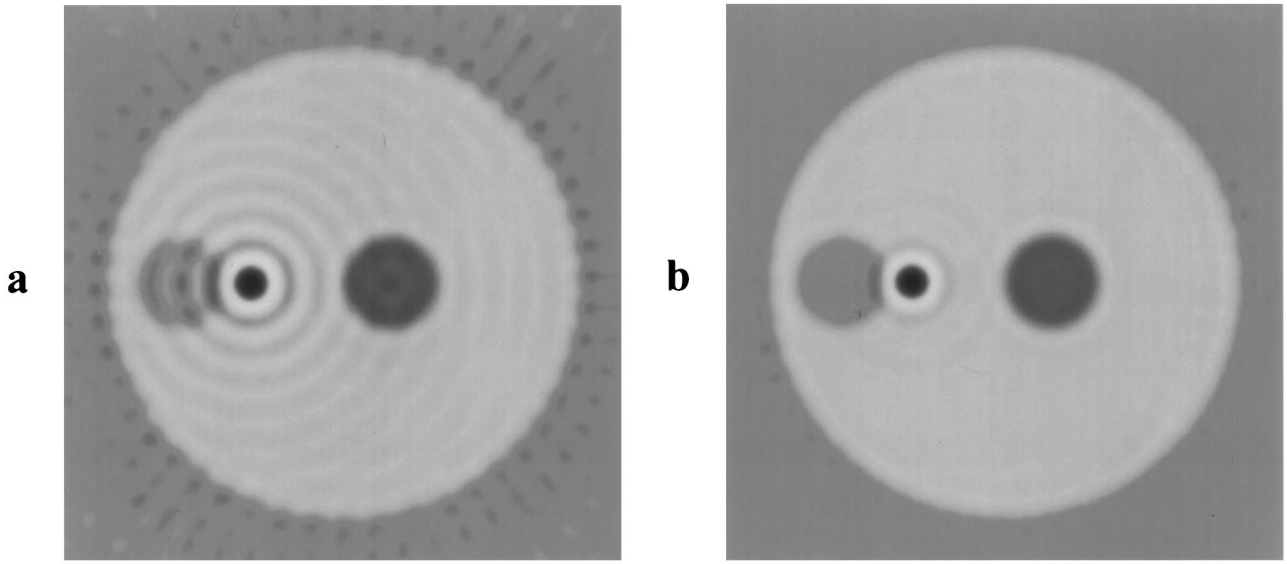


FIG. 6. Eigenfunction images of a test object having background compressibility contrast 0.01, a cystic (water-filled) region, a pointlike scatterer, and an internal cylinder. The images are shown on a logarithmic scale with a 40 dB dynamic range. (a) Real part of inversion obtained for a wave number such that $ka=25$. (b) Analogous image obtained using five wavenumbers such that $20<ka<30$.

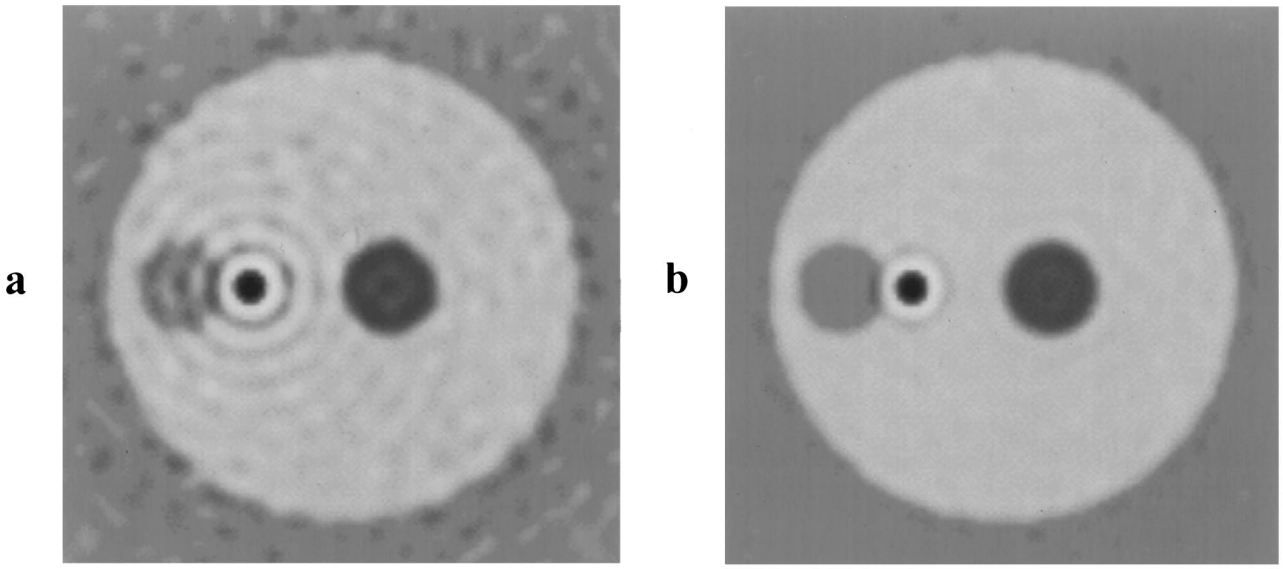


FIG. 7. Images of test object obtained using noised data with 6 dB signal-to-noise ratio. The images are shown on a logarithmic scale with a 40 dB dynamic range. (a) Real part of inversion obtained for a wave number such that $ka=25$. (b) Analogous image obtained using five wave numbers such that $20<ka<30$.

be useful for media such as biological tissue when absorption is finite but small.

A disadvantage of the inverse scattering method as currently implemented is that nonlinear inversion requires an accurate specification of the background acoustic field. This disadvantage is not unique to the eigenfunction method, but is a common feature of most current inversion methods. Recent theoretically exact methods, while not limited in this manner,[31,32] have not yet been implemented numerically.

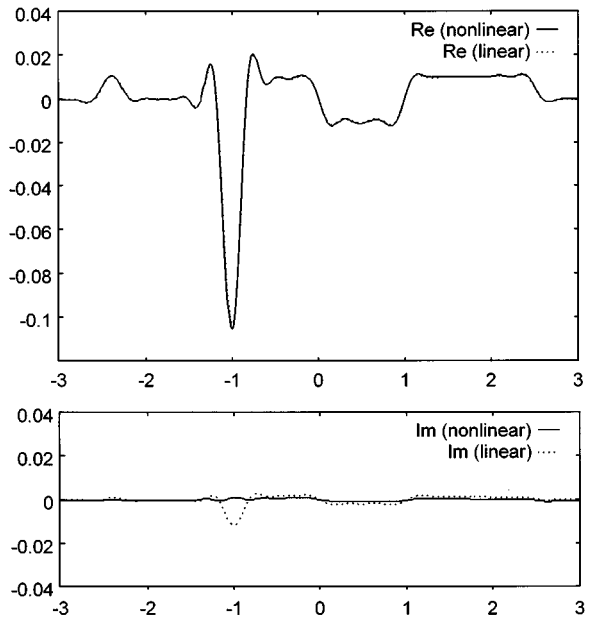


FIG. 8. Cross sections of real and imaginary parts of test object reconstructions using data from five wave numbers. Linear reconstructions were obtained using retransmitted fields in a homogeneous medium, while nonlinear reconstructions were obtained using retransmitted fields in a homogeneous medium and in a background cylinder of compressibility contrast 0.01. (a) Cross sections of real parts. (b) Cross sections of imaginary parts.

The general, nonlinear version of the eigenfunction imaging approach, as defined in Eqs. (24) and (44) and illustrated in Fig. 8, has obvious application to iterative reconstruction of unknown inhomogeneities. In such reconstructions, the total pressure field can be estimated at each iteration from a numerical solution of the direct scattering problem for the currently estimated inhomogeneity, and an eigenfunction inversion can be performed using this pressure field. This procedure can then be repeated to obtain improved estimates of the scattering potential until convergence is achieved.

V. CONCLUSION

A method for focusing and imaging using scattered fields has been presented. The method outlined here makes use of the physical properties of scattering operators by using their eigenfunctions as incident-wave patterns. The eigenfunctions have been shown to provide optimal focusing on pointlike and distributed scattering objects.

The inverse scattering scheme presented exploits the focusing properties of eigenfunctions as well as recent analytic results to obtain a robust and efficient means of quantitatively reconstructing unknown scattering media. This new method has a complexity dependent only on the size and complexity of the scattering medium. Particular cases of the method provide improved implementations of eigenfunction focusing and filtered backpropagation. The method can be implemented for any background medium for which the total acoustic pressure field can be estimated.

Another particular case of the presented method results in a nonlinear inverse scattering formula that yields a solution for the scattering potential q in terms of retransmitted fields of eigenfunctions in the scattering medium and in the background medium. This formula has been demonstrated to yield improvement in accuracy and resolution over Born inversion.

The ideas reported here are expected to be useful in further studies of inverse scattering, adaptive focusing, and ultrasonic imaging. The eigenfunctions of the scattering operator A, whether determined by iterative retransmission or by numerical diagonalization, may be used to focus through inhomogeneous media and to determine optimal incident-wave patterns for inverse scattering experiments. Also, the products of fields of eigenfunctions are expected to form a useful basis for expansion of unknown scattering media in iterative reconstruction algorithms.

ACKNOWLEDGMENTS

The authors thank Peter Monk for releasing his scattering problem solver under the Gnu General Public License and for assistance in the use of his code. Helpful discussions with Dong-Lai Liu, Petri Ola, and Fadil Santosa are acknowledged with pleasure. Funding for this work was provided by the F. V. Hunt Fellowship of the Acoustical Society of America, U.S. Army Grant No. DAMD17-94-J-4384, the University of Rochester Diagnostic Ultrasound Research Laboratory Industrial Associates, NIH Grants No. DK 45533 and No. HL 150855, and ONR Grant No. N00014-91-1107.

[1] D. Isaacson, ''Distinguishability of conductivities by electric current computed tomography,'' IEEE Trans. Med. Imaging **MI-5**, 91–95 (1986).
[2] D. G. Gisser, D. Isaacson, and J. C. Newell, ''Electric current computed tomography and eigenvalues,'' SIAM (Soc. Ind. Appl. Math.) J. Appl. Math. **50**, 1623–1634 (1990).
[3] E. Isaacson and H. S. Keller, *Analysis of Numerical Methods* (Wiley, New York, 1966), pp. 147–159.
[4] G. S. Agarwal, A. T. Friberg, and E. Wolf, ''Scattering theory of distortion correction by phase conjugation,'' J. Opt. Soc. Am. **73**, 529–538 (1983).
[5] R. Mittra and T. M. Habashy, ''Theory of wave-front distortion by phase conjugation,'' J. Opt. Soc. Am. A **1**, 1103–1109 (1984).
[6] O. Ikeda, ''An image reconstruction algorithm using phase conjugation for diffraction-limited imaging in an inhomogeneous medium,'' J. Acoust. Soc. Am. **85**, 1602–1606 (1989).
[7] D. R. Jackson and D. R. Dowling, ''Phase conjugation in underwater acoustics,'' J. Acoust. Soc. Am. **89**, 171–181 (1991).
[8] M. Fink, ''Time-reversal of ultrasonic fields—Part I: Basic principles,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **39**, 555–567 (1992).
[9] D. Cassereau and M. Fink, ''Time-reversal of ultrasonic fields—Part III: Theory of the closed time-reversal cavity,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **39**, 579–592 (1992).
[10] C. Prada, J. -L. Thomas, and M. Fink, ''The iterative time-reversal process: analysis of the convergence,'' J. Acoust. Soc. Am. **97**, 62–71 (1995).
[11] C. Prada, S. Manneville, D. Spoliansky, and M. Fink, ''Decomposition of the time reversal operator: Detection and selective focusing on two scatterers,'' J. Acoust. Soc. Am. **99**, 2067–2076 (1996).
[12] K. Chadan and P. C. Sabatier, *Inverse Problems in Quantum Scattering Theory* (Springer-Verlag, New York, 1989).
[13] A. J. Devaney, ''A filtered backpropagation algorithm for diffraction tomography,'' Ultrason. Imaging **4**, 336–350 (1982).
[14] A. J. Devaney and G. Beylkin, ''Diffraction tomography using arbitrary transmitter and receiver surfaces,'' Ultrason. Imaging **6**, 181–193 (1984).
[15] A. Witten, J. Tuggle, and R. C. Waag, ''A practical approach to ultrasonic imaging using diffraction tomography,'' J. Acoust. Soc. Am. **83**, 1645–1652 (1988).
[16] J. R. Taylor, *Scattering Theory* (Wiley, New York, 1972).
[17] P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (McGraw-Hill, New York, 1968), Chap. 8.
[18] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953), Vol. I.
[19] M. Reed and B. Simon, *Methods of Mathematical Physics* (Academic, San Diego, 1980), Vol. III.
[20] G. N. Watson, *A Treatise on the Theory of Bessel Functions* (Macmillan, New York, 1944), 2nd ed., p. 359.
[21] D. C. Dobson and F. Santosa, ''Resolution and stability analysis of an inverse problem in electrical impedance tomography: dependence on the input current patterns,'' SIAM (Soc. Ind. Appl. Math.) J. Appl. Math. **54**, 1542–1560 (1994).
[22] T. M. Habashy, R. W. Groom, and B. R. Spies, ''Beyond the Born and Rytov approximations: a nonlinear approach to electromagnetic scattering,'' J. Geophys. Res. **98**, 1759–1775 (1993).
[23] A. Kirsch and P. Monk, ''An analysis of the coupling of finite element and Nyström methods in acoustic scattering,'' IMA J. Num. Anal. **14**, 523–544 (1994).
[24] T. 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).
[25] D. 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).
[26] W. C. Chew and G. P. Otto, ''Microwave imaging of multiple conducting cylinders using local shape functions,'' IEEE Microwave Guided Wave Lett. **2**, 284–286 (1992).
[27] H. 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).
[28] A. Derode, P. Roux, and M. Fink, ''Acoustic time-reversal through high-order multiple scattering,'' *1995 IEEE Ultrasonics Symposium Proceedings* (IEEE, Piscataway, NJ, 1995), Vol. 2, pp. 1091–1094.
[29] D. Colton and R. Kress, ''Eigenvalues of the far field operator and inverse scattering theory,'' SIAM J. Math. Anal. **26**, 601–615 (1995). One of the anonymous reviewers is thanked for drawing the authors' attention to Refs. 29 and 30.
[30] D. Colton and R. Kress, ''Eigenvalues of the far field operator for the Helmholtz equation in an absorbing medium,'' SIAM (Soc. Ind. Appl. Math.) J. Appl. Math. **55**, 1724–1735 (1995).
[31] A. I. Nachman, ''Reconstructions from boundary measurements,'' Ann. Math. **128**, 531–576 (1988).
[32] A. I. Nachman, ''Global uniqueness for a two-dimensional inverse boundary problem,'' Ann. Math. **143**, 71–96 (1996).

Acoustic axisymmetric radiation and scattering from bodies of revolution using the internal source density and Fourier methods

Peter R. Stepanishen
Department of Ocean Engineering, University of Rhode Island, Kingston, Rhode Island 02881-0814

(Received 8 May 1996; revised 5 February 1997; accepted 13 February 1997)

A brief review is presented of a previously developed mean-square error method to solve the Neumann boundary-value problem corresponding to acoustic axisymmetric harmonic radiation and rigid body scattering from bodies of revolution of arbitrary shape. The method is based on the use of an internal line monopole source distribution along the axis. In contrast to earlier work, the pressures are simply expressed as line integrals of the source distributions which in the far field reduce to Fourier transform relationships. Exact expressions for the source strength distributions for a sphere are developed using the Fourier transform relationships. The source strength distributions involve spatial derivatives of Dirac delta functions and thus have a vanishingly small region of support about the center of the sphere. For cylindrically shaped bodies with large aspect ratios the spatial characteristics of the source strength distributions are shown to be more closely matched to the normal velocity distribution. The use of a discrete series of Dirac delta functions to represent the continuous line source distributions is presented for spherical radiation and scattering problems. [S0001-4966(97)00407-4]

PACS numbers: 43.20.Fn, 43.20.Rz, 43.20.Tb [JEG]

INTRODUCTION

Acoustic harmonic radiation and rigid body scattering radiation and scattering from rigid bodies may be formulated as classical boundary-value problems in physics. Bowman *et al.*[1] have compiled a comprehensive summary of numerical results for electromagnetic and acoustic scattering by bodies of various shapes for which the wave equation is separable. Extensive numerical results for three-dimensional harmonic problems are presented for point-source and plane-wave excitations. In particular, numerical results are presented for spherical, prolate spheroidal, and oblate spheroidal bodies.

There are numerous classical analytical methods[1–4] which have been developed to address acoustic harmonic radiation and rigid body scattering problems. These methods include low-frequency expansions and high-frequency asymptotic techniques, e.g., classical ray theory and the geometrical theory of diffraction. However, integral equation methods[5–7] are the present choice for addressing the general three-dimensional acoustic harmonic radiation and scattering problem at low and mid frequencies for arbitrary geometries.

In the present paper an alternative to the integral equation approach is presented to address acoustic three-dimensional harmonic radiation and rigid body scattering problems from bodies of revolution. The approach is based on a previously developed mean-square (MS) error method[8–10] which uses an internal line monopole source distribution along the axis of symmetry. In contrast to the earlier method, pressures are simply obtained here from line integrals of the source distributions which in the far field reduce to Fourier transform relationships. A similar approach and development for the acoustic two-dimensional harmonic radiation and rigid body scattering problem from cylindrical bodies with a plane of symmetry has also been recently developed.[11]

In contrast to the earlier papers[8–10] in which the MS and internal source density method was utilized to obtain numerical results for specific spherical and prolate spheroidal geometries, closed-form solutions for the monopole line source distributions for selected geometries are developed in the present paper. More specifically, examples are presented to illustrate the general spatial characteristics of the source strength distributions for bodies of revolution with widely varying aspect ratios. These examples provide insight into the nature of the frequency and spatial dependence of the source strength distributions for spheres and long thin cylinders.

Finally, a discrete series approximation using spatial Dirac delta functions for the monopole line-source distribution is introduced and several classical examples involving acoustic radiation from a sphere are then solved in closed form. These examples provide additional insight into the accuracy of the internal source density method using such approximations to evaluate the source density distributions and the associated pressure fields. Limited numerical results for acoustic plane-wave scattering by a rigid sphere are presented to illustrate the accuracy of the basic approach to evaluate the surface pressure and the general spatial characteristics of the source strength distributions. Numerical results for more general spheroidal bodies have been previously presented.[8–10]

I. THEORY

Consider a smooth body of revolution which is in contact with an external ideal fluid as shown in Fig. 1. Since the acoustic axisymmetric harmonic radiation and/or rigid body

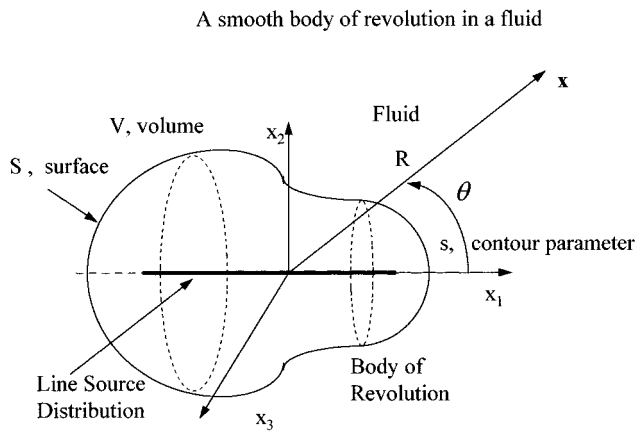


FIG. 1. A smooth body of revolution in a fluid.

scattering problem is of interest, the total pressure field in the external fluid is to be determined for an incident harmonic pressure field and/or the harmonic normal velocity on the surface of the body which are specified to be axisymmetric about x_1.

It is convenient to express the total pressure field for the acoustic harmonic radiation/scattering problem of interest as

$$p_t(\mathbf{x}) = p_i(\mathbf{x}) + p(\mathbf{x}), \tag{1}$$

where the total pressure field $p_t(\mathbf{x})$ is decomposed into the incident harmonic pressure field $p_i(\mathbf{x})$ and the pressure field $p(\mathbf{x})$ which accounts for the rigid body scattering and/or the radiation due to the specified normal velocity. For the radiation problem, the normal velocity of the surface S, $v_r(\mathbf{x}_s)$, is a specified function of s, a contour parameter.

Now $p(\mathbf{x})$ can be posed as the solution of the following Neumann boundary value problem:

$$\nabla^2 p(\mathbf{x}) + k^2 p(\mathbf{x}) = 0, \quad \mathbf{x} \in V, \tag{2a}$$

$$v(s) = \frac{-1}{jk\rho_0 c_0} \frac{\partial p(\mathbf{x}_s)}{\partial n} \quad \mathbf{x}_s(s) \in S$$

$$= \frac{1}{jk\rho_0 c_0} \frac{\partial p_i(\mathbf{x}_s)}{\partial n} + v_r(\mathbf{x}_s), \tag{2b}$$

$$\lim_{r \to \infty} r\left[\frac{\partial p}{\partial r} + jkp\right] = 0, \tag{2c}$$

where k is the acoustic wave number, n denotes the exterior normal on the surface S, i.e., into the fluid volume V, and ρ_0 and c_0 are the density and acoustic wave speed of the fluid. Equation (2b) is the specified Neumann boundary condition and Eq. (2c) is the Sommerfeld radiation condition which is required to ensure uniqueness of the solution. The pressure field $p(\mathbf{x})$ is considered here to be the superposition of the radiated and scattered fields and interaction with the incident field is neglected.

In order to determine the pressure field $p(\mathbf{x})$ in the external fluid consider now a line distribution of unknown monopole sources along the axis of symmetry of the body as shown in Fig. 1. The associated pressure field can be expressed as the following line integral along the axis of symmetry of the body:

$$p(\mathbf{x}) = jk\rho_0 c_0 \int g(\mathbf{x}|\mathbf{x}_1)\sigma(x_1)dx_1, \tag{3}$$

where $\sigma(x_1)$ denotes the unknown line-source density (source strength per unit length), and $g(\mathbf{x}|\mathbf{x}_1)$ is the three-dimensional free-space Green's function, i.e.,

$$g(\mathbf{x}|\mathbf{x}_1) = \frac{e^{-jkr}}{4\pi r}, \tag{4}$$

where $r = |\mathbf{x} - \mathbf{x}_1|$. It then follows from the linearized momentum equation that the associated normal velocity at the surface can be expressed as

$$v_m(s) = -\int \frac{\partial g(\mathbf{x}_s|\mathbf{x}_1)}{\partial n} \sigma(x_1)dx_1, \tag{5}$$

where s and x_1 denote positions along the contour and the axis of symmetry, respectively.

A previously developed MS or mean-square error criteria is now used to determine the unknown source strengths of the internal monopole source distribution.[8–10] More specifically, J is introduced as follows to measure the error between the specified normal surface velocity $v(s)$ and that due to the monopole source density distribution:

$$J = \int |v(s) - v_m(s)|^2 \, dS. \tag{6}$$

The procedure to minimize J thus requires the solution of the following integral equation of the first kind for the unknown source distribution

$$v(s) = -\int \frac{\partial g(\mathbf{x}_s|\mathbf{z})}{\partial n} \sigma(z)dz, \tag{7}$$

where $v(s)$ is known and $\mathbf{z} = z\hat{\mathbf{x}}_1$ is used to denote position along the x_1 axis.

Closed-form solutions for $\sigma(z)$ in Eq. (7) for spherical and cylindrical geometries are presented in the following section; however, in general, closed-form solutions for the unknown line source density $\sigma(z)$ cannot be obtained for an arbitrarily shaped body. A general result for the internal monopole source distribution is, however, readily apparent for bodies of revolution which are doubly symmetric, i.e., surfaces whose cross sections are symmetric about both the x_1 and x_2 axes. If the normal axisymmetric velocity on the surface S is symmetric or antisymmetric about x_2, the monopole distribution clearly exhibits the same symmetry as a function of $z = x_1$.

For the general case of a body of arbitrary shape numerical approximation methods must thus be employed. Such methods lead to the following generic representation for $\sigma(z)$:

$$\sigma(z) = \sum_{i=1}^{N} h_i(z)\sigma_i, \tag{8}$$

where the $h_i(z)$ may be a suitable set of local basis or interpolation functions which are defined within a subinterval about z_i. As a specific example of a local basis function, $h_i(z)$ could be represented as

$$h_i(z) = \delta(z - z_i), \tag{9}$$

where $\delta(z)$ is a Dirac delta function. The local basis functions in Eq. (9) result in a useful discrete approximation to the continuous distribution $\sigma(z)$ as noted later.

It now follows from Eqs. (7) and (8) that $v(s)$ can be represented as

$$v(s)=\sum_{i=1}^{N} K(s|z_i)\sigma_i, \tag{10a}$$

where the $K(s|z_i)$ are simply expressed as

$$K(s|z_i)=-\int \frac{\partial g(\mathbf{x}_s|\mathbf{z})}{\partial n} h_i(z)\, dz. \tag{10b}$$

For an $M>N$ point collocation on the surface S it then follows from Eq. (10) that

$$v(s_m)=\sum_{i=1}^{N} K(s_m|z_i)\sigma_i, \quad m=1,\ldots,M \tag{11}$$

which can also be expressed in matrix form as

$$[K]\boldsymbol{\sigma}=\mathbf{v}, \tag{12}$$

where $[K]$ is an $M\times N$ matrix with $M>N$. SVD methods[12] can then be used to express $\boldsymbol{\sigma}$, the best approximation in the least-squares sense, as

$$\boldsymbol{\sigma}=[K]^{\dagger}\mathbf{v}, \tag{13}$$

where $[K]^{\dagger}$ denotes the pseudo-inverse of $[K]$.

Now $\sigma(z)$ can be simply evaluated from Eq. (8) and the associated pressure field can then be simply expressed as follows from Eq. (3) for all $\mathbf{x}$:

$$p(\mathbf{x})=\sum_{i=1}^{N} G(\mathbf{x}|z_i)\sigma_i, \tag{14a}$$

where the $G(\mathbf{x}|z_i)$ are simply defined as

$$G(\mathbf{x}|z_i)=jk\rho_0 c_0 \int g(\mathbf{x}|\mathbf{z})h_i(z)\, dz. \tag{14b}$$

For spatial points in the far field of the body of interest, the free-space Green's function can be approximated as follows:

$$g(\mathbf{x}|\mathbf{z})\sim \frac{e^{jkR}}{4\pi R} e^{jk\cos\theta z}, \tag{15}$$

where (R,θ) denotes the spherical coordinates of the point of interest. The far-field pressure can then be expressed as

$$p(R,\theta)\sim jk\rho_0 c_0 \frac{e^{jkR}}{4\pi R}\overline{\sigma}(k\cos\theta), \tag{16}$$

where $\sigma(k_z)$ is a one-dimensional Fourier transform of the source distribution, i.e.,

$$\overline{\sigma}(k_z)=\int_{-\infty}^{\infty}\sigma(z)e^{jk_z z}\, dz. \tag{17}$$

Finally, for the case where the local basis functions in Eq. (9) are used to approximate $\sigma(z)$ it is apparent that

$$\sigma(z)=\sum_{i=1}^{N}\delta(z-z_i)\sigma_i \tag{18}$$

and $\sigma(k_z)$ can then be expressed as

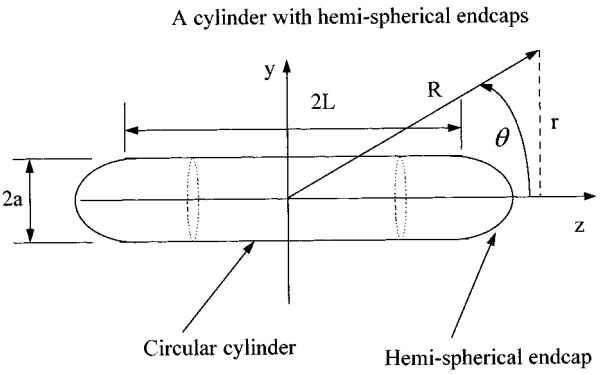


FIG. 2. A cylinder with hemispherical end caps.

$$\overline{\sigma}(k_z)=\sum_{i=1}^{N} e^{jk_z x_i}\sigma_i. \tag{19}$$

For equispaced z_i it is thus apparent that the directional characteristics of the pressure field can be readily obtained via the use of fast Fourier transform (FFT) algorithms.

II. SPECIAL CASES

In order to provide some insight into the characteristics of the internal source density distributions for axisymmetric bodies, the geometry shown in Fig. 2 consisting of a cylinder of length L and radius a with spherical end caps is now considered. The aspect ratio for the indicated cylinder with hemispherical end caps is $L/a+1$. Although the general case of interest may be addressed using the numerical approach based on singular value decomposition (SVD), the limiting cases of L/a zero and infinite are addressed here. These cases correspond to spherical and cylindrical geometries, respectively. Although edge effects are not present in the latter case, the analysis provides insight into the use of the MS method for axisymmetric bodies with large aspect ratios.

The acoustic radiation and scattering from a sphere are classical boundary-value problems which are readily solved using the method of separation of variables.[13] Since closed-form solutions for these problems can be readily obtained, such solutions can be used to obtain insight into the nature of the source distributions for bodies with aspect ratios close to unity.

In general the axisymmetric normal velocity distribution on a sphere with a radius a can be expressed as the following series:

$$v(a,\theta)=\sum_{n=0}^{\infty} V_n P_n(\cos\theta), \tag{20}$$

where $P_n()$ denotes the nth-order Legendre polynomial, θ is the polar angle relative to the axis of symmetry, and the V_n are assumed to be known. The associated axisymmetric pressure field in the fluid can then be expressed as a function of the spherical coordinates (R,θ)

$$p(R,\theta)=-j\rho_0 c_0\sum_{n=0}^{\infty} V_n \frac{h_n^{(2)}(kR)}{h_n^{(2)\prime}(ka)} P_n(\cos\theta), \tag{21}$$

where $h_n^{(2)}()$ denotes the spherical Hankel function of the second kind. It is thus apparent that the associated source density distribution $\sigma(z)$ can be expressed as

$$\sigma(z)=\sum_{n=0}^{\infty}\sigma_n(z), \tag{22}$$

where $\sigma_n(z)$ is a line source density which is associated with the nth-order Legendre polynomial.

It is sufficient to consider only a single azimuthal harmonic of arbitrary order to obtain insight into the nature of the internal source distribution. After noting the asymptotic form of $h_n(kR)$ it is readily apparent from Eq. (21) that the far-field pressure for the nth azimuthal mode can be expressed as

$$p_n(R,\theta)\sim j^n\rho_0 c_0 V_n\,\frac{e^{jkR}}{4\pi kR}\,\frac{P_n(\cos\,\theta)}{h_n^{(2)'}(ka)}. \tag{23}$$

The following Fourier transform or integral equation of the first kind is then obtained for $\sigma_n(z)$ by equating the results in Eqs. (16), (17), and (23)

$$\int_{-\infty}^{\infty}\sigma_n(z)e^{jk\,\cos\,\theta z}dz=\sigma_n P_n(\cos\,\theta), \tag{24a}$$

where σ_n is defined as

$$\sigma_n=j^{n-1}\,\frac{4\pi a^2 V_n}{(ka)^2 h_n^{(2)'}(ka)}. \tag{24b}$$

In order to determine $\sigma_n(z)$ using Eq. (24), it is first noted that $P_n(\cos\,\theta)$ can be expressed as an nth-order polynomial in $\cos\,\theta$, i.e.,

$$P_n(\cos\,\theta)=\sum_{i=0}^{n}b_i(\cos\,\theta)^i, \tag{25}$$

where the b_i are constants where $b_i=0$ for i odd(even) and n even(odd). It then follows that

$$\int_{-\infty}^{\infty}\sigma_n(z)e^{jk\,\cos\,\theta z}\,dz=\sigma_n\sum_{i=0}^{n}b_i\,\cos^i\,\theta. \tag{26}$$

After noting the following property[14] of the Dirac delta function $\delta(z)$

$$\int_{-\infty}^{\infty}\delta^{(p)}(z)e^{jk\,\cos\,\theta z}\,dz=(-jk\,\cos\,\theta)^p, \tag{27}$$

where $\delta^{(p)}(z)$ denotes the pth derivative, it then follows that

$$\sigma_n(z)=\sigma_n\sum_{i=0}^{n}b_i\,\frac{\delta^{(i)}(z)}{(-jk)^i}. \tag{28}$$

In light of the exact solution of the internal source distribution in Eq. (28), several observations regarding the use of the MS method for spherical bodies are readily apparent. Since $\delta^{(I)}(z)$ is an even(odd) generalized function for I even(odd), and it was noted that $b_i=0$ for I odd(even) and n even(odd), then $\sigma_{m,n}(z)$ is an even(odd) function for n even(odd). It then follows from earlier discussion that the pressure is zero on the x_2 axis for n odd. This is of course an expected result since the associated field is proportional to $P_n(\cos\,\theta)$ which is zero for $\theta=\pi/2$ and n odd.

The most striking features about the exact solution of the internal source distributions in Eq. (28) are the singular behavior and the limited region of support of the source distribution about the origin due to nature of the Dirac delta function and its derivatives. Clearly as n increases the region of support for any discrete numerical approximation of $\sigma_n(z)$ will increase, the peak amplitudes will increase, and the resultant approximation will become more oscillatory. Also, as ka increases the maximum number of n for a specified problem will also generally increase.

It is thus to be expected that high-frequency problems will be more formidable to address using the MS method. In contrast to the use of the normal equations to solve the least mean-square error problem for small ka, the present analysis indicates that it becomes increasingly necessary to utilize the SVD methods noted earlier for large ka. These conclusions are consistent with the observations noted during recent numerical studies conducted with the method.

Consider now the limiting case of the geometry shown in Fig. 2 where $L/a\rightarrow\infty$ and end effects of the cylinder are thus omitted. The axisymmetric normal or radial velocity is denoted in cylindrical coordinates (r,z,ϕ) as

$$v(a,z,0)=V_0e^{j\beta z},\quad r=a, \tag{29}$$

where z, which replaces s used earlier, is measured from the origin $z=0$ and β is a constant. It is apparent from Eq. (7) that $\sigma(z)$ must satisfy the following integral equation of the first kind:

$$V_0e^{j\beta z}=-\int_{-\infty}^{\infty}\frac{\partial g(\mathbf{x}_s|\mathbf{z})}{\partial r}\,\sigma(z)dz,\quad r=a. \tag{30}$$

As a result of the periodicity of the normal velocity distribution in z, the internal monopole source distribution may be represented as

$$\sigma(z)=\sigma_0e^{j\beta z}, \tag{31}$$

where σ_0 is to be determined. It then follows from Eqs. (30) and (31) that

$$V_0=-\sigma_0\,\frac{\partial I}{\partial r},\quad y=a, \tag{32a}$$

where, due to the symmetric nature of the integrand, I is expressed as

$$I=\frac{1}{4\pi}\int_{-\infty}^{\infty}\frac{e^{-jk\sqrt{r^2+z^2}}}{\sqrt{r^2+z^2}}\,e^{j\alpha z}\,dz\,. \tag{32b}$$

Since[15]

$$\frac{-j}{4}\,H_0(\sqrt{k^2-\alpha^2}r)=\frac{1}{4\pi}\int_{-\infty}^{\infty}\frac{e^{-jk\sqrt{r^2+z^2}}}{\sqrt{r^2+z^2}}\,e^{j\alpha z}\,dz, \tag{33}$$

where $H_0()$ denotes the zeroth-order Hankel function of the second kind, then σ_0 may be expressed as

$$\sigma_0 = \frac{V_0}{\sqrt{k^2-\beta^2}H_0'(\sqrt{k^2-\beta^2}a)}. \tag{34}$$

If $\sigma(z)$ is substituted into Eq. (6a) and the resultant integral evaluated using Eq. (33), the following expression is then obtained for the pressure field:

$$p(r,z,0) = \frac{k\rho_0 c_0 V_0}{\sqrt{k^2-\beta^2}}\, e^{j\beta z}\, \frac{H_0(\sqrt{k^2-\beta^2}r)}{H_0'(\sqrt{k^2-\beta^2}a)}. \tag{35}$$

The expression in Eq. (35) clearly satisfies the boundary conditions and the Helmholtz equation and is thus the solution to the boundary value problem of interest. Of course for this problem the same solution is obtained more directly via the method of separation of variables.

The solution for $\sigma(z)$ defined in Eqs. (31) and (34) is noted to exhibit significantly different characteristics than the monopole source density distribution for the sphere. In contrast to the sphere, the region of support is not confined to the origin but is in fact the entire axis of symmetry. In addition, the spatial dependence of the source density for the ''large aspect cylindrical body of revolution'' matches the spatial variation of the normal velocity on the surface which is in contrast to that of the sphere.

It is now noted that β is a free parameter in the solution for the ''large aspect cylinder.'' If V_0 is considered to be a function of β, more general spatially bounded normal velocity distributions may be represented via an integration over β, i.e.,

$$v(a,z,0) = \frac{1}{2\pi}\int_{-\infty}^{\infty} V_0(\beta)e^{-j\beta z}\, d\beta. \tag{36a}$$

It is then apparent that $V_0(\beta)$ is the Fourier transform of the spatially bounded normal velocity distribution $v(a,z,0)$, i.e.,

$$V_0(\beta) = \int_{-\infty}^{\infty} v(a,z,0)e^{j\beta z}\, dz. \tag{36b}$$

The associated $\sigma(z)$ for the spatially bounded normal velocity distributions of interest can also be obtained by integration over β

$$\sigma_m(z) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{V_0(\beta)}{\sqrt{k^2-\beta^2}H_0'(\sqrt{k^2-\beta^2}a)}\, e^{-j\beta z}\, d\beta. \tag{37}$$

It then follows from Eq. (37) that the source distribution will, in general, be different than the normal velocity distribution. It is noted, however, that the spatially bounded normal velocity and source strength distributions become more similar as the spectral bandwidth of the velocity decreases.

After performing an integration over β in Eq. (35) the following expression for the associated pressure is obtained:

$$p(r,z,0) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{k\rho_0 c_0 V_0(\beta)}{\sqrt{k^2-\beta^2}}\, \frac{H_0(\sqrt{k^2-\beta^2}r)}{H_0'(\sqrt{k^2-\beta^2}a)}\, e^{j\beta z}\, d\beta. \tag{38}$$

This latter expression is equivalent to a previous result used to study the forward and backward projection of acoustic fields from finite cylindrical vibrators with infinite rigid cylindrical extensions.[16]

III. A DISCRETE APPROXIMATION FOR $\sigma(z)$

Consider again the special case of a sphere of radius a. The use of a discrete approximation for the continuous source distributions of the sphere are now investigated. For the sake of simplicity the source distributions in Eq. (22) for all cases are represented using Eqs. (8) and (9) as

$$\sigma_n(z) = \sum_{m=-M}^{M} \sigma_n^{(m)}\delta(z-m\Delta), \tag{39}$$

where $M\Delta < a$ is required in order that the source distributions be contained within the sphere and the constants $\sigma_n^{(m)}$ are now to be determined.

The case of $n=0$ is almost trivial and is first addressed. In light of Eqs. (24b), (25), and (28) it is apparent that

$$\sigma_0(z) = \sigma_0\delta(z). \tag{40}$$

It then follows from Eq. (14) that the pressure field can be expressed as

$$p(\mathbf{x}) = jk\rho_0 c_0 \sigma_0 g(\mathbf{x}|0). \tag{41}$$

After using Eq. (24b) for $n=0$ and noting that $g(\mathbf{x}|0) = -jkh_0^{(2)}(kR)/4\pi$ it then follows that

$$p(\mathbf{x}) = -j\rho_0 c_0 V_0\, \frac{h_0^{(2)}(kR)}{h_0^{(2)'}(ka)}, \tag{42}$$

which is the exact solution.

The $n=1$ case can also be simply addressed since it is noted from Eq. (25) that $b_1 = 1$ and $b_i = 0$ for $i \neq 1$. From Eq. (28) it is then apparent that

$$\sigma_1(z) = \sigma_1\, \frac{\delta^{(1)}(z)}{(-jk)}, \tag{43}$$

which can be approximated using a central difference method as

$$\sigma_1(z) = \frac{\sigma_1}{jk\Delta}\, \delta(z-\Delta) - \frac{\sigma_1}{jk\Delta}\, \delta(z+\Delta). \tag{44}$$

It is now obvious that $\sigma_1^{(1)} = -\sigma_1^{(-1)} = \sigma_1\Delta/jk$ and $\sigma_n^{(m)} = 0$ for other m and n. It then follows from Eq. (14) that the pressure field for this case can be expressed as

$$p(\mathbf{x}) = \frac{\rho_0 c_0 \sigma_1}{\Delta}\, [g(\mathbf{x}|-\mathbf{z}) - g(\mathbf{x}|+\mathbf{z})], \tag{45}$$

where $\mathbf{z} = \Delta\hat{x}_1$. In the limit as $\Delta \to 0$ it is apparent that $p(\mathbf{x})$ can be expressed as

$$p(\mathbf{x}) = \rho_0 c_0 \sigma_1\, \left.\frac{\partial g(\mathbf{x}|\mathbf{z})}{\partial z}\right|_{\mathbf{z}=0}. \tag{46}$$

It then follows from the limiting case of the finite difference approximation for the derivative noted earlier that $p(\mathbf{x})$ can simply be expressed as

$$p(\mathbf{x}) = -j\rho_0 c_0 V_1\, \frac{h_1^{(2)}(kR)}{h_1^{(2)'}(ka)}\, P_1(\cos\,\theta), \tag{47}$$

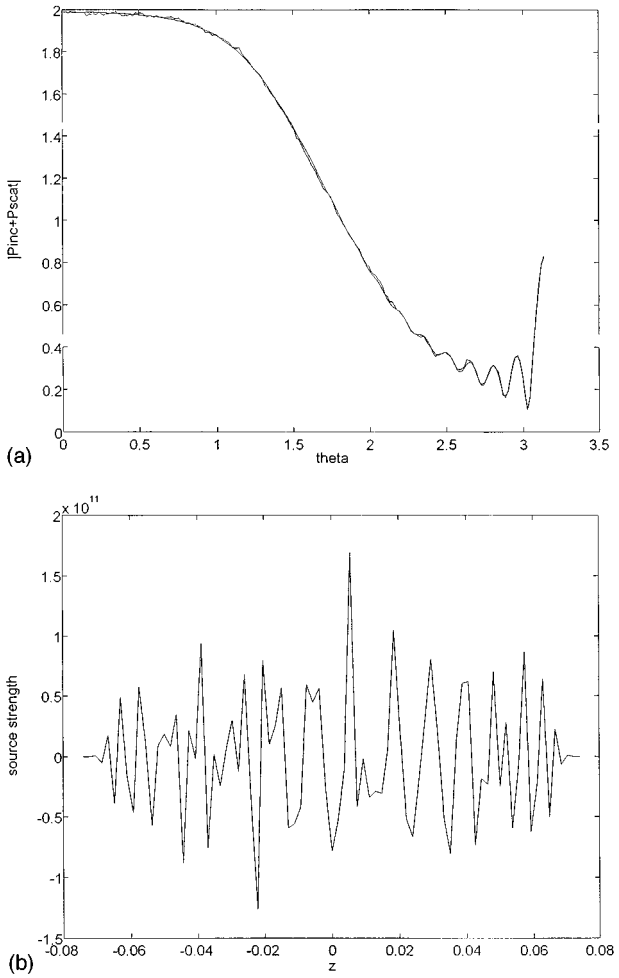


FIG. 3. Surface pressure on a rigid sphere versus θ (in radians) and the associated monopole source strength represented as a continuous function of position. (a) Surface pressure. (b) Monopole source strength.

which is the exact solution for the field.

Several comments regarding the use of the discrete approximation in Eq. (39) are now apparent from Eq. (45). It is noted that: The source strengths coefficients for the $n=1$ azimuthal mode are proportional to Δ^{-1}, the region of support for the $n=1$ distribution is 2Δ, and successive coefficients are of opposite sign for the $n=1$ case. It can be simply shown by repeating the procedure for higher n that: The source strengths coefficients for the nth azimuthal mode are proportional to Δ^{-n}, the region of support for the nth azimuthal mode increases with n, and successive coefficients for a fixed n are of opposite sign.

In order to illustrate the nature of the source distribution for a problem of interest, consider now plane-wave scattering from a rigid sphere. A closed-form solution is readily obtained for the scattered pressure in the form of Eq. (21) and the total pressure field can then be evaluated using Eq. (1). In general, as ka increases an increasing number of terms is required in order for the modal series to converge.

Numerical results are presented for the surface pressure and associated source density distribution for $ka=20$ in Fig. 3(a) and (b), respectively. In contrast to the preceding examples where a modal decomposition was used, the source distribution is the total distribution $\sigma(z)$ which was obtained via the use of Eq. (12). It is apparent from Fig. 3(a) that the surface pressure obtained via the LMS method is in excellent agreement with the exact analytical modal solution in both the illuminated and shadow regions of the sphere. The results for the source density distribution in Fig. 3(b) are concentrated into a spatial region of $|z|<0.2a$. The large amplitude oscillatory behavior is to be expected in light of the preceding discussion. Similar numerical results have also been obtained in earlier studies using the LMS method for the acoustic axisymmetric harmonic radiation and rigid body scattering from spherical bodies.

IV. SUMMARY AND CONCLUSIONS

A previously developed MS error method[8–10] is used to solve the acoustic harmonic radiation and rigid body scattering problem for bodies of revolution. Both problems are simply combined into a single Neumann boundary-value problem which is solved via the use of an internal line monopole distribution along the axis of symmetry. The MS method leads to an inhomogeneous integral equation of the first kind for the unknown line monopole distribution. Field pressures are simply obtained from line integrals of the source distributions which in the far field reduce to Fourier transform relationships.

In contrast to earlier investigations of the MS method in which extensive numerical studies were performed, the focus of the present study was to determine closed-form analytical solutions to specific problems. These problems were selected to represent two broad classes of bodies of interest, i.e., bodies with aspect ratios near 1 and bodies with large aspect ratio. The former class of problems encompass spheres while the latter class encompass long thin cylinders.

Exact expressions for the source strength distributions corresponding to acoustic two-dimensional harmonic radiation and rigid body scattering from spheres are developed using asymptotic and Fourier transform relationships. The resulting source strength distributions involve spatial derivatives of Dirac delta functions and thus have a vanishingly small region of support about the center of the sphere. The spatial characteristics of the distributions indicate that problems involving spheres provide formidable test problems for the general MS approach presented here.

In order to provide insight into the nature of the line-source distributions for cylindrical shaped bodies of revolution with large aspect ratios, the spatial characteristics of the source strength distributions were investigated for the case of a long thin cylinder with hemispherical end caps. Ignoring. end effects, it is apparent that the spatial characteristics of the source strength distributions are different than those of the normal velocity distribution. The differences decrease as the spectral purity of the velocity increases, i.e., as the velocity becomes more narrow band in wave-number space. The resulting expression for the pressure field was also noted to agree with a previously developed expression which was developed using separation of variables and a wave-number transform technique.

Since discrete numerical methods are in general required to use the MS method to address the axisymmetric three-

dimensional harmonic radiation and scattering problems from an arbitrary body of revolution, the effects of using a discrete series approximation for the source distributions are briefly addressed. The discrete series approximation is used for the spherical problem and the resulting expressions for the source strength distributions are noted to exhibit the same characteristics as those which were observed using SVD methods during previous numerical studies. A single numerical result for a scattering problem from a sphere is also presented to further illustrate the behavior of the source distribution and the pressure field.

In concluding, it appears that the proposed MS method is a viable approach for addressing the axisymmetric three-dimensional acoustic harmonic radiation and scattering problems from smooth bodies of revolution. It is easily implemented and circumvents many of the numerical problems which arise in using integral equation methods. Although the focus in the present paper is on the acoustic problem, the approach is obviously applicable to a wide range of physical problems which are described mathematically by the reduced wave equation and Neumann boundary conditions.

Finally it is noted the basic MS or internal source density and SVD approach has been recently generalized to address the nonaxisymmetric problem.[17] It is not surprising that such an extension is possible since the analogous approach for two-dimensional problems has already been completed and reported.[11,18–20]

[1] J. J. Bowman, T. B. A. Senior, and P. L. E. Uslenghi, *Electromagnetic and Acoustic Scattering by Simple Shapes* (North-Holland, Amsterdam, 1969).

[2] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953).

[3] I. Stakgold, *Boundary Value Problems in Mathematical Physics* (Macmillan, New York, 1967).

[4] D. G. Crighton, A. P. Dowling, J. E. Ffowcs Williams, M. Heckl, and F. G. Leppington, in *Modern Methods in Analytical Acoustics*, Lecture Notes (Springer-Verlag, London, 1992).

[5] L. H. Chen and D. G. Schweikert, ''Sound radiation from an arbitrary body,'' J. Acoust. Soc. Am. **35**, 1626–1632 (1963).

[6] H. A. Schenck, ''Improved integral formulation for acoustic radiation problems,'' J. Acoust. Soc. Am. **44**, 41–58 (1967).

[7] A. J. Burton and G. F. Miller, ''The application of integral equation methods to the numerical solution of some exterior boundary-value problems,'' Proc. R. Soc. London, Ser. A **323**, 201–210 (1971).

[8] P. R. Stepanishen and H. W. Chen, ''Surface pressure and harmonic loading on shells of revolution using an internal source density method,'' J. Acoust. Soc. Am. **92**, 2248–2259 (1992).

[9] P. R. Stepanishen and H. W. Chen, ''Acoustic harmonic radiation and scattering from shells of revolution using finite element and internal source density methods,'' J. Acoust. Soc. Am. **92**, 2248–2259 (1992).

[10] P. R. Stepanishen and H. W. Chen, ''Acoustic time-dependent loading on elastic shells of revolution using the internal source density and singular value decomposition method,'' J. Acoust. Soc. Am. **99**, 1913–1923 (1996).

[11] P. R. Stepanishen, ''Acoustic two-dimensional radiation and scattering from cylinders using source density, SVD and Fourier methods, J. Sound Vib. **199**(3), 1–17 (1997).

[12] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes* (Cambridge U. P., Cambridge, England, 1986).

[13] M. C. Junger and D. Feit, *Sound Sructures and their Interaction* (MIT, Cambridge, MA, 1972).

[14] A. Papoulis, *Signal Analysis* (McGraw-Hill, New York, 1977).

[15] I. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series, and Products*, edited by A. Jeffreys (Academic, London, 1980).

[16] P. R. Stepanishen and H. W. Chen, ''Near field pressures and surface intensity for cylindrical vibrators,'' J. Acoust. Soc. Am. **76**, 942–948 (1984).

[17] P. R. Stepanishen, ''A generalized internal source density method for the forward and backward projection of harmonic pressure fields from complex bodies,'' J. Acoust. Soc. Am. **101**, 3270–3277 (1997).

[18] P. R. Stepanishen and S. Ramakrishna, ''Acoustic radiation from cylinders with a plane of symmetry using internal multipole source distributions. I,'' J. Acoust. Soc. Am. **93**, 658–672 (1993).

[19] S. Ramakrishna and P. R. Stepanishen, ''Acoustic scattering from cylinders with a plane of symmetry using internal multipole source distributions. II,'' J. Acoust. Soc. Am. **93**, 673–682 (1993).

[20] P. R. Stepanishen and S. Ramakrishna, ''Acoustic impedances and impulse responses for elliptical cylinders using internal source density and SVD methods,'' J. Sound Vib. **176**, 49–68 (1994).

Space–time far-field representation of Green's functions for cross-plane shear waves in general transversely isotropic media

Martin Spies
Fraunhofer-Institut Zerstörungsfreie Prüfverfahren (IZFP), Universität des Saarlandes, Geb. 37, 66123 Saarbrücken, Germany

(Received 10 April 1996; revised 14 February 1997; accepted 19 February 1997)

For transversely isotropic materials like fiber composites, but also unidirectionally grain-structured austenitic steels, the elastodynamic dyadic and triadic Green's functions for the pure shear (SH) wave mode are given in the space–time domain. The first stage of the derivation is performed in a way similar to the derivation of these functions in the form of their 2D space–time spectral representations [M. Spies, J. Acoust. Soc. Am. **96**, 1144–1157 (1994)]. In the present study, the inverse Fourier integrations are performed using a far-field approximation. The resulting expressions appear in a coordinate-free form and contain the orientation of the material's fiber axis as an additional parameter. Thus the results are particularly useful for extension to the case of layered material. Numerical evaluation is performed for a unidirectional composite material for several orientations. Since the application of the SH-wave type bears considerable potential in nondestructive testing, the results presented in this article may find some application within this framework. © *1997 Acoustical Society of America.* [S0001-4966(97)00108-2]

PACS numbers: 43.20.Gp, 43.20.Jr, 43.35.Cg, 43.35.Zc [JEG]

INTRODUCTION

The application of structural materials like composites/composite laminates and columnar-grained steels raises the demand for appropriate testing and quality inspection. However, due to the high degree of anisotropy, conventional testing methods based on elastic wave propagation fail in general when applied to these media. Therefore, considerable interest is dedicated to studying wave propagation in elastically anisotropic materials, as well as scope to the wave scattering problem. In this respect, Green's dyadic and triadic functions—the kernel of the fundamental mathematical formulation of Huygens' principle by Pao and Varatharajulu[1]—are of considerable importance. Their significance in the development of wave propagation theory for isotropic media is well-known; corresponding analytical formulations for anisotropic materials are scarce and mostly restrictive and approximative in character.[2–4] To account for the generality required in many problems, a number of recent studies have dealt with numerical Green's function models.[5–8]

For the case of transversely isotropic (TI) symmetry, which is characteristic for unidirectional fiber composites and extruded metal-matrix composites, but also for ideally fiber-textured columnar-grained steels, analytical representations for Green's dyadic and triadic functions have recently been given.[9] Since the material's spatial orientation has been included as an additional parameter, which is most convenient in view of layered structures, the functions have been obtained in the form of their 2D space–time spectral representations to be evaluated using fast Fourier transformation. The functions are given by the sum of three terms, each one generating the different patterns of wave propagation related to the quasi-pressure, quasi-shear, and pure shear wave inherent to transversely isotropic materials. The analytical determination of these tensor Green's functions in the space–time domain has not been possible due to the complex analytical expressions characterizing the quasi-wave terms.

However, as shown in this study, far-field representations of Green's dyadic and triadic functions can be obtained in the space–time domain for the pure shear wave propagating in TI media. This wave is usually called the SH wave, while the coupled waves are called the quasi-longitudinal (qP) and quasi-transverse (qSV) wave. In order to emphasize reference to the planes of symmetry, Helbig[10] used ''cross-plane S'' for SH, ''in-plane waves'' collectively for the two coupled waves, and ''in-plane S'' for the qSV wave. Although SH will be used throughout this article for historical and notational reasons, the designation ''cross-plane S'' seems, from the physical point of view, to be more correct. Since the application of this wave type bears considerable potential in nondestructive testing,[11–14] the results presented in this article may find some application within this framework.[15,16]

In Sec. I, fundamental relationships of plane-wave theory for transversely isotropic media with arbitrary spatial orientation (TIA media),[17] as relevant to this article, are summarized. Section II deals with with the construction of Green's dyadic (or displacement tensor[1]) function using an approach similar to the one taken in Ref. 9. On this basis, the ''cross-plane S'' or SH part of the dyad is derived in space–time domain in Sec. III, making use of a far-field approximation. Finally, the analytical results, given in coordinate-free representation, are evaluated and discussed in Sec. IV for a TI composite material. The derivation of Green's triadic (or stress tensor[1]) function, which is in principle straightforward, is given in the Appendices.

I. PLANE-WAVE PROPAGATION IN TIA MEDIA

The equation of motion for the displacement vector $\underline{\mathbf{u}}$ in a homogeneous anisotropic solid reads[1]

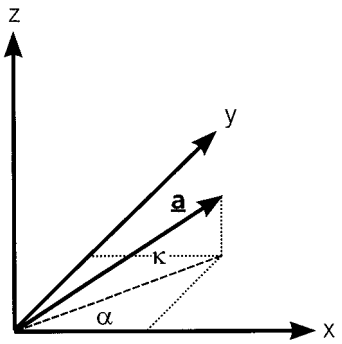


FIG. 1. The fiber direction $\underline{\mathbf{a}}$ is variable in 3D space, specified by α, the angle between its projection onto the x-y plane and the x axis, and κ, the angle between its x-y projection and itself.

$$(\underline{\nabla}\cdot\underline{\underline{\mathbf{C}}}\cdot\underline{\nabla})\cdot\underline{\mathbf{u}}(\underline{\mathbf{R}},\omega)+\varrho\,\omega^2\underline{\mathbf{u}}(\underline{\mathbf{R}},\omega)=-\underline{\mathbf{f}}(\underline{\mathbf{R}},\omega), \tag{1}$$

where $\underline{\nabla}$ is the gradient vector, ϱ is the material density, $\underline{\mathbf{f}}$ accounts for the volume force density, and ω denotes the circular frequency, if time dependence $\sim e^{-j\omega t}$ is assumed. The elastic stiffness tensor $\underline{\underline{\mathbf{C}}}$ is for the general transversely isotropic case given by

$$\begin{aligned}\underline{\underline{\mathbf{C}}}(\underline{\mathbf{a}})=&\lambda_\perp\underline{\underline{\mathbf{II}}}+\mu_\perp[(\underline{\underline{\mathbf{II}}})^{1324}+(\underline{\underline{\mathbf{II}}})^{1342}]+(\nu-\lambda_\perp)[\underline{\mathbf{Iaa}}\\&+\underline{\mathbf{aaI}}]+(\mu_\parallel-\mu_\perp)[(\underline{\mathbf{Iaa}})^{1324}+(\underline{\mathbf{aaI}})^{1324}\\&+(\underline{\mathbf{Iaa}})^{1342}+(\underline{\mathbf{aaI}})^{1342}]+[\lambda_\perp+2\mu_\perp+\lambda_\parallel+2\mu_\parallel\\&-2(\nu+2\mu_\parallel)]\underline{\mathbf{aaaa}},\end{aligned} \tag{2}$$

where unit vector $\underline{\mathbf{a}}$ indicates the orientation of the material's rotational symmetry axis, which will in the following be referred to as the ''fiber direction.'' $\underline{\underline{\mathbf{I}}}$ is the dyadic idemfactor, which can be decomposed, using dyadic products of the Cartesian unit vectors, according to

$$\underline{\underline{\mathbf{I}}}=\underline{\mathbf{e}}_x\underline{\mathbf{e}}_x+\underline{\mathbf{e}}_y\underline{\mathbf{e}}_y+\underline{\mathbf{e}}_z\underline{\mathbf{e}}_z=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}; \tag{3}$$

accordingly, the tetrad $\underline{\underline{\mathbf{II}}}$ is the dyadic product of $\underline{\underline{\mathbf{I}}}$ with itself. The upper indicial notation indicates transposition of elements in the tetrads; $\lambda_\parallel+2\mu_\parallel$, $\lambda_\perp+2\mu_\perp$, $\mu_\perp$, $\mu_\parallel$, and ν[8] are the elastic constants corresponding to C_{11}, C_{33}, C_{44}, C_{66}, and C_{13}, respectively, in Voigt's abbreviated notation, when the fiber direction is parallel to the x direction (Fig. 1).

The plane harmonic wave solutions of (1) with the stiffness tensor given by Eq. (2), i.e., the solutions to the homogeneous ($\underline{\mathbf{f}}=\underline{\mathbf{0}}$) equation of motion, are in the form

$$\underline{\mathbf{u}}_\alpha(\underline{\mathbf{R}},\omega)=U\hat{\underline{\mathbf{u}}}_\alpha\,\exp[jK_\alpha\hat{\underline{\mathbf{K}}}\cdot\underline{\mathbf{R}}], \tag{4}$$

where $\hat{\underline{K}}$ is the propagation direction. The wave numbers K_α are given by the dispersion relations as[17]

$$K_{SH}^2=\varrho\,\omega^2/[\mu_\perp+(\mu_\parallel-\mu_\perp)(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{K}}})^2], \tag{5}$$

$$K_{qSV,qP}^2=\varrho\,\omega^2(B\pm(B^2-4A)^{1/2})/(2A), \tag{6}$$

where

$$\begin{aligned}A=&\mu_\parallel(\lambda_\perp+2\mu_\perp)+(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{K}}})^2(\lambda_\parallel(\lambda_\perp+2\mu_\perp)-\nu(\nu\\&+2\mu_\parallel))+(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{K}}})^4[(\lambda_\parallel+\mu_\parallel)(\mu_\parallel-(\lambda_\perp+2\mu_\perp))\\&+(\nu+\mu_\parallel)^2],\end{aligned} \tag{7}$$

$$B=\mu_\parallel+\lambda_\perp+2\mu_\perp+(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{K}}})^2(\lambda_\parallel+2\mu_\parallel-(\lambda_\perp+2\mu_\perp)). \tag{8}$$

The discrimination into the pure transverse wave (SH: shear horizontal) and the two quasi-waves (qSV: quasi-shear vertical; qP: quasi-pressure) becomes apparent when the respective polarizations $\hat{\underline{\mathbf{u}}}_\alpha$ are considered, which are given by

$$\hat{\underline{\mathbf{u}}}_{SH}=\mathcal{N}_{SH}(\underline{\mathbf{a}}\times\hat{\underline{\mathbf{K}}}), \tag{9}$$

$$\hat{\underline{\mathbf{u}}}_{qSV,qP}=\mathcal{N}_{qSV,qP}(F_1^{qSV,qP}\hat{\underline{\mathbf{K}}}+F_2^{qSV,qP}\underline{\mathbf{a}}). \tag{10}$$

The abbreviated quantities can be found in Ref. 17. Since the SH polarization is always perpendicular to the $\underline{\mathbf{a}}$-$\hat{\underline{\mathbf{K}}}$ plane, i.e., in general not horizontally polarized, whereas the quasi-wave polarizations lie within this plane, Helbig's designations as *cross-plane S* wave and *in-plane* waves, respectively, are less misleading.[10]

II. CONSTRUCTION OF GREEN'S DYADIC FUNCTION

Green's dyadic function $\underline{\underline{\mathbf{G}}}$ is defined by the differential equation[1]

$$(\underline{\nabla}\cdot\underline{\underline{\mathbf{C}}}\cdot\underline{\nabla})\cdot\underline{\underline{\mathbf{G}}}(\underline{\mathbf{R}},\omega)+\varrho\,\omega^2\underline{\underline{\mathbf{G}}}(\underline{\mathbf{R}},\omega)=-\underline{\underline{\mathbf{I}}}\delta(\underline{\mathbf{R}}), \tag{11}$$

where δ is Dirac's three-dimensional delta function. Application of spatial Fourier transform (FT) with respect to $\underline{\mathbf{R}}$ in terms of

$$\tilde{\underline{\underline{\mathbf{G}}}}(\underline{\mathbf{K}},\omega)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\underline{\underline{\mathbf{G}}}(\underline{\mathbf{R}},\omega)e^{-j\underline{\mathbf{K}}\cdot\underline{\mathbf{R}}}\,d^3\underline{\mathbf{R}} \tag{12}$$

results in

$$\tilde{\underline{\underline{\mathbf{W}}}}(\underline{\mathbf{K}},\omega)\cdot\tilde{\underline{\underline{\mathbf{G}}}}(\underline{\mathbf{K}},\omega)=\underline{\underline{\mathbf{I}}}. \tag{13}$$

This facilitates the construction of the dyadic Green's function which is obtained according to

$$\underline{\underline{\mathbf{G}}}(\underline{\mathbf{R}},\omega)=\frac{1}{(2\pi)^3}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\tilde{\underline{\underline{\mathbf{W}}}}^{-1}(\underline{\mathbf{K}},\omega)e^{j\underline{\mathbf{K}}\cdot\underline{\mathbf{R}}}\,d^3\underline{\mathbf{K}}. \tag{14}$$

Here the wave matrix

$$\tilde{\underline{\underline{\mathbf{W}}}}(\underline{\mathbf{K}},\omega)=\underline{\mathbf{K}}\cdot\underline{\underline{\mathbf{C}}}\cdot\underline{\mathbf{K}}-\varrho\,\omega^2\underline{\underline{\mathbf{I}}} \tag{15}$$

appears, which is explicitly given by[18]

$$\tilde{\underline{\underline{\mathbf{W}}}}(\underline{\mathbf{K}},\omega)=\alpha\underline{\underline{\mathbf{I}}}+\beta\underline{\mathbf{KK}}+\gamma\underline{\mathbf{aa}}+\epsilon(\underline{\mathbf{Ka}}+\underline{\mathbf{aK}}), \tag{16}$$

where

$$\alpha=\mu_\perp K^2+(\mu_\parallel-\mu_\perp)(\underline{\mathbf{a}}\cdot\underline{\mathbf{K}})^2-\varrho\,\omega^2, \tag{17}$$

$$\beta=\lambda_\perp+\mu_\perp, \tag{18}$$

$$\gamma=(\lambda_\perp+2\mu_\perp+\lambda_\parallel+2\mu_\parallel-2(\nu+2\mu_\parallel))(\underline{\mathbf{a}}\cdot\underline{\mathbf{K}})^2 + (\mu_\parallel-\mu_\perp)K^2, \tag{19}$$

$$\epsilon=((\nu+\mu_\parallel)-(\lambda_\perp+\mu_\perp))(\underline{\mathbf{a}}\cdot\underline{\mathbf{K}}). \tag{20}$$

The integral representation (14) of the dyadic Green's function satisfies the radiation condition and thus constitutes the solution of Eq. (11) that leads to a physically reasonable result.[19,20]

The inverse of $\underline{\underline{\tilde{\mathbf{W}}}}(\underline{\mathbf{K}},\omega)$ is obtained according to

$$\underline{\underline{\tilde{\mathbf{W}}}}^{-1}(\underline{\mathbf{K}},\omega)=\frac{\text{adj }\underline{\underline{\tilde{\mathbf{W}}}}(\underline{\mathbf{K}},\omega)}{\det\underline{\underline{\tilde{\mathbf{W}}}}(\underline{\mathbf{K}},\omega)}, \tag{21}$$

where the adjoint, determined using dyadic algebra, is given by[9,18]

$$\text{adj }\underline{\underline{\tilde{\mathbf{W}}}}(\underline{\mathbf{K}},\omega)=\frac{1}{\alpha}(\det\underline{\underline{\tilde{\mathbf{W}}}})\underline{\underline{\mathbf{I}}}-(\alpha\beta+\beta\gamma-\epsilon^2)\underline{\mathbf{K}}\underline{\mathbf{K}}-[\alpha\gamma+(\beta\gamma-\epsilon^2)K^2]\underline{\mathbf{a}}\underline{\mathbf{a}}-[\alpha\epsilon-(\beta\gamma-\epsilon^2)\times(\underline{\mathbf{a}}\cdot\underline{\mathbf{K}})](\underline{\mathbf{K}}\underline{\mathbf{a}}+\underline{\mathbf{a}}\underline{\mathbf{K}}). \tag{22}$$

The determinant of $\underline{\underline{\tilde{\mathbf{W}}}}$ is given by

$$\det\underline{\underline{\tilde{\mathbf{W}}}}(\underline{\mathbf{K}},\omega)=\alpha[\alpha(\alpha+\beta K^2+\gamma+2\epsilon(\underline{\mathbf{a}}\cdot\underline{\mathbf{K}}))+(\beta\gamma-\epsilon^2)(K^2-(\underline{\mathbf{a}}\cdot\underline{\mathbf{K}})^2)]\equiv\alpha\Lambda, \tag{23}$$

where α characterizes the *SH*-wave propagation, while the quasi-waves are characterized by Λ as can be easily seen from Eqs. (17) to (20).

Thus the inverse of the wave matrix can be written as

$$\underline{\underline{\tilde{\mathbf{W}}}}^{-1}(\underline{\mathbf{K}},\omega)=\frac{1}{\alpha}\underline{\underline{\mathbf{I}}}-\frac{(\beta\gamma-\epsilon^2)}{\alpha\Lambda}[\underline{\mathbf{K}}\underline{\mathbf{K}}+K^2\underline{\mathbf{a}}\underline{\mathbf{a}}-(\underline{\mathbf{a}}\cdot\underline{\mathbf{K}})(\underline{\mathbf{K}}\underline{\mathbf{a}}+\underline{\mathbf{a}}\underline{\mathbf{K}})]-\frac{1}{\Lambda}(\beta\underline{\mathbf{K}}\underline{\mathbf{K}}+\gamma\underline{\mathbf{a}}\underline{\mathbf{a}}+\epsilon(\underline{\mathbf{K}}\underline{\mathbf{a}}+\underline{\mathbf{a}}\underline{\mathbf{K}})). \tag{24}$$

Using this equation, the 3D space–time spectral representation of Green's dyadic function has been derived in Ref. 9. Due to the complex structure of the wave numbers K_{qSV}^2 and K_{qP}^2 [Eq. (6)], the backtransformation of $\underline{\underline{\tilde{\mathbf{G}}}}(\underline{\mathbf{K}},\omega)$ into spatial domain does not succeed, thus a 2D space–time spectral representation to be evaluated via FFT has been obtained.[9,18] However, since the wave number K_{SH}^2 [Eq. (5)] is less complicated in structure, a space–time representation of $\underline{\underline{\mathbf{G}}}$ can be derived for the *SH* part as is done in the next section.

III. DERIVATION OF $\underline{\underline{\mathbf{G}}}_{SH}$ IN THE SPACE–TIME DOMAIN

A. *SH* part of the inverse wave matrix

Performing a partial fracture expansion of Eq. (24) with respect to α and Λ and sorting for the *SH* and quasi-wave terms, respectively, yields the *SH* part according to

$$\underline{\underline{\tilde{\mathbf{W}}}}_{SH}^{-1}(\underline{\mathbf{K}},\omega)=\frac{1}{\alpha}\underline{\underline{\mathbf{I}}}+\frac{\mu_\perp}{\alpha\sigma}[\underline{\mathbf{K}}\underline{\mathbf{K}}+K^2\underline{\mathbf{a}}\underline{\mathbf{a}}-(\underline{\mathbf{a}}\cdot\underline{\mathbf{K}})(\underline{\mathbf{K}}\underline{\mathbf{a}}+\underline{\mathbf{a}}\underline{\mathbf{K}})], \tag{25}$$

where

$$\sigma=\mu_\parallel(\underline{\mathbf{a}}\cdot\underline{\mathbf{K}})^2-\varrho\omega^2. \tag{26}$$

Inserting Eq. (25) into Eq. (14), replacing wave vector $\underline{\mathbf{K}}$ in the numerator by $-j\underline{\nabla}$, and interchanging the order of differentiation and integration, the *SH* part of Green's dyad is obtained as

$$\underline{\underline{\mathbf{G}}}_{SH}(\underline{\mathbf{R}},\omega)=\underline{\underline{\mathbf{I}}}\,I_1(\underline{\mathbf{R}},\omega)-[\underline{\nabla}\underline{\nabla}+\underline{\nabla}^2\underline{\mathbf{a}}\underline{\mathbf{a}}-(\underline{\mathbf{a}}\cdot\underline{\nabla})(\underline{\nabla}\underline{\mathbf{a}}+\underline{\mathbf{a}}\underline{\nabla})]\,I_2(\underline{\mathbf{R}},\omega), \tag{27}$$

where

$$I_1(\underline{\mathbf{R}},\omega)=\frac{1}{(2\pi)^3}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{e^{j\underline{\mathbf{K}}\cdot\underline{\mathbf{R}}}\,d^3\underline{\mathbf{K}}}{\mu_\perp K^2+(\mu_\parallel-\mu_\perp)(\underline{\mathbf{a}}\cdot\underline{\mathbf{K}})^2-\varrho\omega^2}, \tag{28}$$

$$I_2(\underline{\mathbf{R}},\omega)=\frac{1}{(2\pi)^3}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{e^{j\underline{\mathbf{K}}\cdot\underline{\mathbf{R}}}\,d^3\underline{\mathbf{K}}}{(\mu_\perp K^2+(\mu_\parallel-\mu_\perp)(\underline{\mathbf{a}}\cdot\underline{\mathbf{K}})^2-\varrho\omega^2)(\mu_\parallel(\underline{\mathbf{a}}\cdot\underline{\mathbf{K}})^2-\varrho\omega^2)}, \tag{29}$$

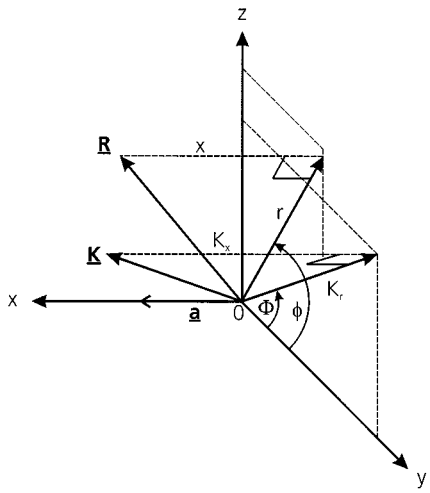


FIG. 2. Coordinate system used in evaluating the integrals I_1 and I_2.

B. Evaluation of integrals

The following considerations are easier to perform using cylindrical coordinates, where for convenience and without loss of generality $\underline{\mathbf{a}}=\underline{\mathbf{e}}_x$ is chosen (Fig. 2). Once the respective results have been obtained, they will be rewritten in coordinate-free form. This approach has been used previously by Chen[20] to evaluate an integral similar to $I_1(\underline{\mathbf{R}},\omega)$. It is

$$\underline{\mathbf{K}}=K_x\ \underline{\mathbf{e}}_x+K_r\ \cos\ \Phi\ \underline{\mathbf{e}}_y+K_r\ \sin\ \Phi\ \underline{\mathbf{e}}_z, \tag{30}$$

$$\underline{\mathbf{R}}=x\ \underline{\mathbf{e}}_x+r\ \cos\ \phi\ \underline{\mathbf{e}}_y+r\ \sin\ \phi\ \underline{\mathbf{e}}_z, \tag{31}$$

$$\underline{\mathbf{K}}\cdot\underline{\mathbf{R}}=xK_x+rK_r\ \cos(\Phi-\phi), \tag{32}$$

the differential volume element is $d^3\underline{\mathbf{K}}=K_r\ dK_r\ dK_x\ d\Phi$, the limits for K_x, K_r, and Φ are from $-\infty$ to ∞, from 0 to ∞, and from 0 to 2π, respectively. Furthermore, α and σ transform into

$$\alpha=\mu_\parallel K_x^2+\mu_\perp K_r^2-\varrho\,\omega^2, \tag{33}$$

$$\sigma=\mu_\parallel K_x^2-\varrho\,\omega^2. \tag{34}$$

Substituting Eq. (32) into Eq. (28) and using Eq. (A1) of Appendix A to perform the Φ integration provides

$$I_1(\underline{\mathbf{R}},\omega)=\frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_0^{\infty}\frac{J_0(rK_r)K_r\ dK_r}{\mu_\perp K_r^2+\mu_\parallel K_x^2-\varrho\,\omega^2}\times e^{jxK_x}\ dK_x, \tag{35}$$

where J_0 is the Bessel function of the first kind of order zero. Using Eq. (A2), the integration over K_r can be performed to yield

$$I_1(\underline{\mathbf{R}},\omega)=\frac{1}{4\pi^2\mu_\perp}\times\int_{-\infty}^{\infty}K_0\left(r\sqrt{\frac{\mu_\parallel}{\mu_\perp}K_x^2-\frac{\varrho\,\omega^2}{\mu_\perp}}\right)e^{jxK_x}\ dK_x, \tag{36}$$

where K_0 is the modified Bessel function of the second kind of order zero. Finally the K_x integration is carried out according to Eq. (A3), and the result is

$$I_1(\underline{\mathbf{R}},\omega)=\frac{1}{4\pi}\frac{\exp(j\omega\sqrt{(\varrho/\mu_\parallel\mu_\perp)\cdot(\mu_\perp x^2+\mu_\parallel r^2)})}{\sqrt{\mu_\perp}\sqrt{\mu_\perp x^2+\mu_\parallel r^2}} \tag{37}$$

or in coordinate-free representation

$$I_1(\underline{\mathbf{R}},\omega)=\frac{1}{4\pi R}\times\frac{\exp(j\omega R\sqrt{\varrho/\mu_\parallel\mu_\perp}\sqrt{\mu_\parallel+(\mu_\perp-\mu_\parallel)(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{R}}})^2})}{\sqrt{\mu_\perp}\sqrt{\mu_\parallel+(\mu_\perp-\mu_\parallel)(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{R}}})^2}}. \tag{38}$$

The evaluation of $I_2(\underline{\mathbf{R}},\omega)$ with respect to the integration over Φ and K_r is performed in the same way as above to give

$$I_2(\underline{\mathbf{R}},\omega)=\frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_0^{\infty}\frac{J_0(rK_r)K_r\ dK_r}{(\mu_\perp K_r^2+\mu_\parallel K_x^2-\varrho\,\omega^2)(\mu_\parallel K_x^2-\varrho\,\omega^2)}\,e^{jxK_x}\ dK_x$$
$$=\frac{1}{4\pi^2\mu_\parallel}\int_{-\infty}^{\infty}\frac{K_0(r\sqrt{\mu_\parallel/\mu_\perp}\sqrt{K_x^2-\varrho\,\omega^2/\mu_\parallel})}{K_x^2-\dfrac{\varrho\,\omega^2}{\mu_\parallel}}\,e^{jxK_x}\ dK_x. \tag{39}$$

Application of Eqs. (A4) and (A5) then leads to

$$I_2(\underline{\mathbf{R}},\omega)=\exp\left(j\omega\sqrt{\frac{\varrho}{\mu_\parallel\mu_\perp}\cdot(\mu_\perp x^2+\mu_\parallel r^2)}\right)\cdot\left\{-\frac{1}{4\pi}\frac{\sqrt{\mu_\perp}}{\varrho\,\omega^2}\frac{\sqrt{\mu_\perp x^2+\mu_\parallel r^2}}{\mu_\parallel r^2}\right.$$
$$\left.+\frac{1}{4\pi}\frac{j}{2\sqrt{\varrho\,\omega^2\mu_\parallel}}\cdot\left(\int_0^{\infty}\frac{e^{-t}\,dt}{(z_1+t)^2}+\int_0^{\infty}\frac{e^{-t}\,dt}{(z_2+t)^2}\right)\right\}, \tag{40}$$

where

$$z_{1,2}=-j\omega\,\frac{\varrho}{\mu_\parallel\mu_\perp}\,(\sqrt{\mu_\perp x^2+\mu_\parallel r^2}\pm\sqrt{\mu_\perp x^2}). \tag{41}$$

In the far field ($x,r\to\infty$), $z_{1,2}$ goes to infinity so that the integrals in (40) vanish, resulting in the coordinate-free representation

$$I_2^{\mathrm{far}}(\underline{\mathbf{R}},\omega)=-\frac{1}{\varrho\,\omega^2}\,\frac{\sqrt{\mu_\perp}\sqrt{\mu_\parallel+(\mu_\perp-\mu_\parallel)(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{R}}})^2}}{\mu_\parallel(1-(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{R}}})^2)}\;\frac{\exp[j\omega R\sqrt{\varrho/\mu_\parallel\mu_\perp}\sqrt{\mu_\parallel+(\mu_\perp-\mu_\parallel)(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{R}}})^2}]}{4\pi R}. \tag{42}$$

C. Explicit coordinate-free form of $\underline{\underline{\mathbf{G}}}_{SH}^{\mathrm{far}}$

Thus resulting from Eq. (27), the far-field representation to be further evaluated is

$$\underline{\underline{\mathbf{G}}}_{SH}^{\mathrm{far}}(\underline{\mathbf{R}},\omega)=\underline{\underline{\mathbf{I}}}\,I_1(\underline{\mathbf{R}},\omega)-[\underline{\nabla}\,\underline{\nabla}+\underline{\nabla}^2\underline{\mathbf{a}}\underline{\mathbf{a}}-(\underline{\mathbf{a}}\cdot\underline{\nabla})(\underline{\nabla}\,\underline{\mathbf{a}}+\underline{\mathbf{a}}\underline{\nabla})]\,I_2^{\mathrm{far}}(\underline{\mathbf{R}},\omega), \tag{43}$$

with I_1 and I_2^{far} given by Eqs. (38) and (42). Performing the $\underline{\nabla}$ operations on I_2^{far} according to

$$\underline{\nabla}R=\hat{\underline{\mathbf{R}}}, \tag{44}$$

$$\underline{\nabla}\,\hat{\underline{\mathbf{R}}}=\frac{1}{R}\,(\underline{\underline{\mathbf{I}}}-\hat{\underline{\mathbf{R}}}\hat{\underline{\mathbf{R}}}) \tag{45}$$

results, after some algebraic manipulations, in

$$[\underline{\nabla}\,\underline{\nabla}+\underline{\nabla}^2\underline{\mathbf{a}}\underline{\mathbf{a}}-(\underline{\mathbf{a}}\cdot\underline{\nabla})(\underline{\nabla}\,\underline{\mathbf{a}}+\underline{\mathbf{a}}\underline{\nabla})]\,I_2^{\mathrm{far}}(\underline{\mathbf{R}},\omega)=\frac{[\hat{\underline{\mathbf{R}}}\hat{\underline{\mathbf{R}}}+\underline{\mathbf{a}}\underline{\mathbf{a}}-(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{R}}})(\hat{\underline{\mathbf{R}}}\underline{\mathbf{a}}+\underline{\mathbf{a}}\hat{\underline{\mathbf{R}}})]}{[1-(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{R}}})^2]}\,I_1(\underline{\mathbf{R}},\omega). \tag{46}$$

Here, relations (B1)–(B4) of Appendix B as well as the condition $R\to\infty$ have been used. Obeying the identity

$$(\underline{\mathbf{a}}\times\hat{\underline{\mathbf{R}}})=[1-(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{R}}})^2]\underline{\underline{\mathbf{I}}}-[\hat{\underline{\mathbf{R}}}\hat{\underline{\mathbf{R}}}+\underline{\mathbf{a}}\underline{\mathbf{a}}-(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{R}}})(\hat{\underline{\mathbf{R}}}\underline{\mathbf{a}}+\underline{\mathbf{a}}\hat{\underline{\mathbf{R}}})], \tag{47}$$

and introducing the unit vector

$$\hat{\underline{\mathbf{u}}}(\hat{\underline{\mathbf{R}}})=\frac{\underline{\mathbf{a}}\times\hat{\underline{\mathbf{R}}}}{[1-(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{R}}})^2]^{1/2}}, \tag{48}$$

then yields

$$\underline{\underline{\mathbf{G}}}_{SH}^{\mathrm{far}}(\underline{\mathbf{R}},\omega)=\sqrt{\frac{\varrho}{\mu_\parallel}}\frac{c_{SH}(\hat{\underline{\mathbf{R}}})}{\mu_\perp}\,\frac{e^{j\omega R/c_{SH}(\hat{\underline{\mathbf{R}}})}}{4\pi R}\,\hat{\underline{\mathbf{u}}}(\hat{\underline{\mathbf{R}}})\hat{\underline{\mathbf{u}}}(\hat{\underline{\mathbf{R}}}), \tag{49}$$

where the modulus of group velocity c_{SH} has been introduced, the derivation of which is described in Appendix C. Performing the inverse Fourier transformation of this expression with respect to ω finally results in

$$\underline{\underline{\mathbf{G}}}_{SH}^{\mathrm{far}}(\underline{\mathbf{R}},t)=\sqrt{\frac{\varrho}{\mu_\parallel}}\frac{c_{SH}(\hat{\underline{\mathbf{R}}})}{\mu_\perp}\,\hat{\underline{\mathbf{u}}}(\hat{\underline{\mathbf{R}}})\hat{\underline{\mathbf{u}}}(\hat{\underline{\mathbf{R}}})\,\frac{\delta(t-R/c_{SH}(\hat{\underline{\mathbf{R}}}))}{4\pi R}, \tag{50}$$

where $\delta(t)$ designates the Dirac-delta function. This is the space–time far-field representation of Green's dyadic function for cross-plane shear waves in TIA media; the corresponding representation of the triadic function is given in Appendix D.

IV. NUMERICAL EVALUATION AND DISCUSSION

Numerical evaluation is performed for a unidirectional graphite-epoxy composite; the relevant material parameters are given by $\mu_\perp=3.4$ GPa, $\mu_\parallel=6.8$ GPa, and $\varrho=1.6$ g/cm^3. Due to the symmetry conditions and the variability of fiber direction $\underline{\mathbf{a}}$ in the results given above, evaluation in the x-z plane is sufficient. In the following, Green's dyadic function is convoluted with an appropriate time function and then plotted for several orientations of fiber direction $\underline{\mathbf{a}}$, which is characterized by angles α and κ as shown in Fig. 1. To account for a realistic experimental pulse as applied in nondestructive testing, a raised cosine (RC2) time function as shown in Fig. 3 has been used according to

$$f_{\mathrm{RC2}}(t)=[1-\cos(\omega t/2)]\cos(\omega t),\quad 0\leqslant t\leqslant 4\pi/\omega. \tag{51}$$

In Figs. 4–6, the modulus of the nonzero components $(\underline{\underline{\mathbf{G}}}_{SH}^{\mathrm{far}})_{ij}$, which, physically, represent the displacement components in j direction due to a concentrated force in i direction, are plotted for a frequency of $\omega=2\pi\cdot 1$ MHz at time $t=5\,\mu$s in linear scale. The represented area is 20×20 mm^2 in the x-z plane, the generating force being located in the center.

In cases where $\alpha=0°$ (Fig. 4) only the component $(\underline{\underline{\mathbf{G}}}_{SH}^{\mathrm{far}})_{yy}$ does not vanish; here a y-directed force excites a y-polarized cross-plane shear wave at the highest amplitude. A variation of κ leads to a respective rotation of the wave fronts. For $\kappa=0°$ fixed, as also displayed in Fig. 4, a variation of α leads to changes in the wave patterns with vanishing contributions in x direction and no excitation at all for $\alpha=90°$. Correspondingly, as shown in Figs. 5 and 6, the other components increase with the characteristic patterns of wave excitation which can, qualitatively speaking, be easily understood according to the nature of the cross-plane SH waves.

The application of the space–time far-field representation of Green's dyadic function for SH waves [Eq. (50)] is

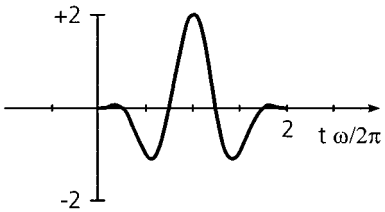


FIG. 3. Raised cosine (RC2) time function used to account for a realistic experimental pulse.

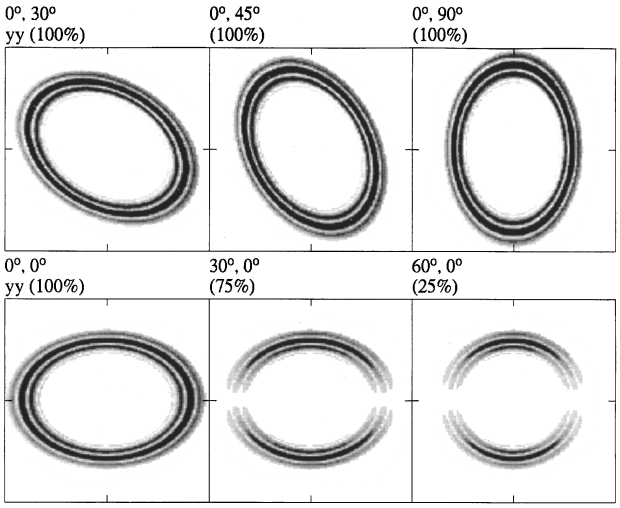


FIG. 4. Component $(\underline{\underline{\mathbf{G}}}^{\mathrm{far}}_{SH})_{yy}$ for $\alpha=0°$ and different values of κ (first row), and for $\kappa=0°$ and variable α (second row). The values in parentheses indicate the respective maximum amplitudes in percentages of the absolute maximum in the case $\alpha=\kappa=0°$.

not restricted to homogeneous TIA media since the orientation of the rotational symmetry axis has been retained as an additional parameter. Thus the results presented in this study are particularly useful in the case of layered structures when emphasis is on *SH*-wave propagation. This is the case, e.g., for two mutually orthogonal layers of composite material with *SH* waves polarized perpendicular to the fibers and obliquely incident on the interface. The resulting reflected waves are also of *SH* type, having the same polarization as the incident wave. Particularly useful for this consideration is the algorithm to numerically evaluate the (plane-wave) reflection and transmission coefficients for the case of two arbitrarily oriented transversely isotropic layers presented previously.[17]

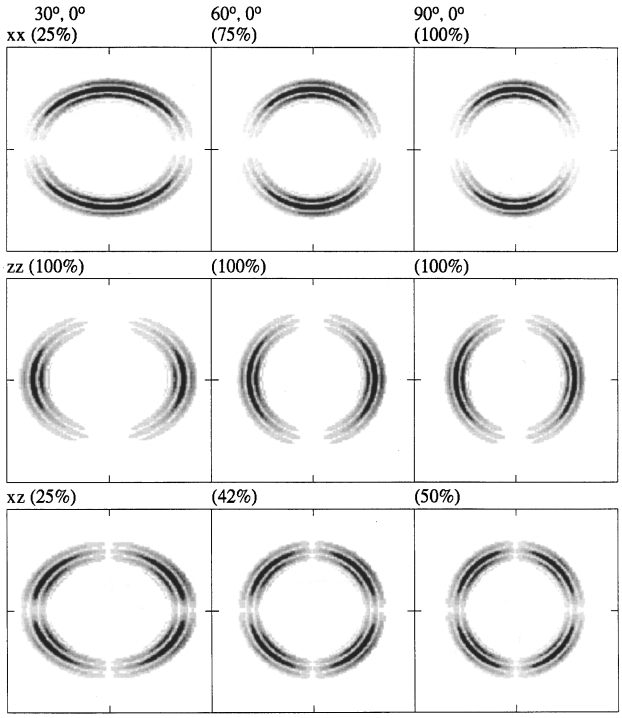


FIG. 5. Components $(\underline{\underline{\mathbf{G}}}^{\mathrm{far}}_{SH})_{xx,zz,xz}$ for $\kappa=0°$ and variable α. The values in parentheses indicate the respective maximum amplitudes in percentages of the absolute maximum in the case $\alpha=\kappa=0°$.

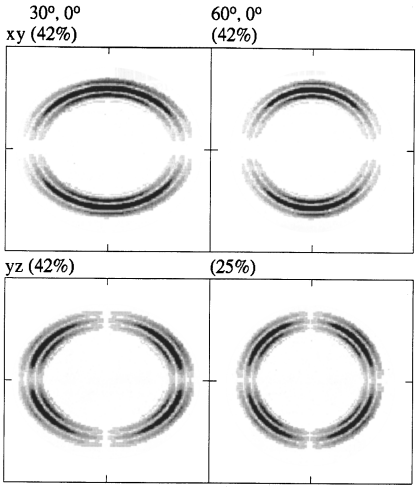


FIG. 6. Components $(\underline{\underline{\mathbf{G}}}^{\mathrm{far}}_{SH})_{xy,yz}$ for $\kappa=0°$ and variable α. The values in parentheses indicate the respective maximum amplitudes in percentages of the absolute maximum in the case $\alpha=\kappa=0°$.

It should be noted that the appearance of group velocity in the exponential function in Eq. (49) is independent of the far-field approximation, as can be seen from Eq. (40), while in the classical paper by Buchwald,[21] group velocity, expressed by the corresponding $\underline{\mathbf{K}}$ vectors, appears through the applied asymptotic integral evaluation. It should also be mentioned that the neglection of the integrals in Eq. (40) does not affect the accuracy of the far-field approximation. Performing the ∇ operations on I_2 instead of I_2^{far} [Eq. (43)] and retaining resulting integrals with terms of order $1/(z_{1,2}+t)^2$ leads to an additional term in Eq. (49), which, however, vanishes when the inverse Fourier transformation with respect to ω is performed. The direct functional relationship between the modulus of group velocity and the spatial coordinates given in the Appendices has proven to be particularly useful for modeling of *SH*-wave propagation at low computation times.[15,16] Finally, it is worth mentioning that the representation given in Eq. (49) leads to the same result as the one given by Ben-Menahem and Gibson[22] for poroelastic inhomogeneous anisotropic media when applied to the homogeneous transversely isotropic case with fiber direction fixed parallel to a coordinate axis.

ACKNOWLEDGMENTS

The author would like to express his thanks to Professor M. Kröning, director of IZFP, Saarbrücken, Germany, and to Professor W. Arnold, head of IZFP's Physical Basics Department, for their interest in and support of the subject. This study was sponsored by the German Department of Education and Research (BMBF), which is gratefully acknowledged.

APPENDIX A. INTEGRAL FORMULAS

The evaluation of the integrals $I_1(\underline{\mathbf{R}},\omega)$ and $I_2(\underline{\mathbf{R}},\omega)$ is performed using the following integral formulas. With J_0

designating the Bessel function of first kind of order zero, K_0 designating the modified Bessel function of the second kind of order zero and E_i being the exponential-integral function[23] it is:

$$\int_0^{2\pi} e^{jx\cos(\Phi-\phi)}d\Phi = 2\pi J_0(x), \tag{A1}$$

$$\int_0^{\infty} \frac{J_0(ax)}{x^2+b^2}\, x\, dx = K_0(ab), \quad a>0, \quad \mathrm{Re}\ b>0, \tag{A2}$$

$$\int_{-\infty}^{\infty} K_0(a\sqrt{x^2-b^2})e^{jcx}\, dx = \frac{\pi}{\sqrt{a^2+c^2}}\exp(jb\sqrt{a^2+c^2}), \tag{A3}$$

$$\int_0^{\infty} \frac{K_0(b\sqrt{a^2+x^2})}{a^2+x^2}\, e^{jcx}\, dx = -\frac{\pi}{4a}\,[e^{-ac}E_i(-d_2) + e^{ac}E_i(-d_1)], \tag{A4}$$

where

$$d_{1,2} = a\sqrt{b^2+c^2} \pm ac,$$

$$E_i(-x) = e^{-x}\left[-\frac{1}{x} + \int_0^{\infty}\frac{e^{-t}}{(x+t)^2}\, dt\right], \quad x>0. \tag{A5}$$

Equation (A3) can be found in Ref. 24, while the other relationships are taken from Ref. 23.

APPENDIX B. SPATIAL DERIVATIVES

In the evaluation of the $\underline{\nabla}$ operations in Eq. (43) the following relationships are used ($D \equiv \mu_\parallel + (\mu_\perp - \mu_\parallel)(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{R}}})^2$, $C \equiv \omega\sqrt{\varrho/\mu_\parallel\mu_\perp}$):

$$\underline{\nabla}(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{R}}}) = \frac{1}{R}\,(\underline{\mathbf{a}} - (\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{R}}})\hat{\underline{\mathbf{R}}}), \tag{B1}$$

$$\underline{\nabla}(R\sqrt{D}) = (\mu_\parallel\hat{\underline{\mathbf{R}}} + (\mu_\perp - \mu_\parallel)(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{R}}})\underline{\mathbf{a}})/\sqrt{D}, \tag{B2}$$

$$\underline{\nabla}\left(\frac{D}{\mu_\parallel[1-(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{R}}})^2]}\right) = 2\frac{\mu_\perp}{\mu_\parallel}\,\frac{(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{R}}})}{R[1-(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{R}}})^2]^2} \times(\underline{\mathbf{a}} - (\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{R}}})\hat{\underline{\mathbf{R}}}), \tag{B3}$$

$$\underline{\nabla}\left(\frac{e^{jCR\sqrt{D}}}{R\sqrt{D}}\right) = \frac{jCR\sqrt{D}-1}{RD}\,\frac{e^{jCR\sqrt{D}}}{R\sqrt{D}}\,(\mu_\parallel\hat{\underline{\mathbf{R}}} + (\mu_\perp-\mu_\parallel) \times(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{R}}})\underline{\mathbf{a}}). \tag{B4}$$

APPENDIX C. DERIVATION OF GROUP VELOCITY AS SPATIAL FUNCTION

In lossless media, the velocity of energy transport is given by the group velocity, which has, for the materials under concern, been obtained previously.[17] For *SH* waves, group velocity can be written as

$$\underline{\mathbf{c}}_{SH} = (\varrho v_{SH})^{-1}[\mu_\perp\hat{\underline{\mathbf{K}}} + (\mu_\parallel-\mu_\perp)(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{K}}})\underline{\mathbf{a}}], \tag{C1}$$

where $v_{SH} = \omega/K_{SH}$ designates the phase velocity. Since the (energy) contribution to the wave field at an observation point $\underline{\mathbf{R}}$ is characterized by group velocity, it is for the respective unit vectors

$$\hat{\underline{\mathbf{R}}} = \hat{\underline{\mathbf{c}}}_{SH}, \quad |\hat{\underline{\mathbf{R}}}| = |\hat{\underline{\mathbf{c}}}_{SH}| = 1. \tag{C2}$$

Thus it is

$$\hat{\underline{\mathbf{R}}} = \frac{\mu_\perp\hat{\underline{\mathbf{K}}} + (\mu_\parallel-\mu_\perp)(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{K}}})\underline{\mathbf{a}}}{\sqrt{\mu_\perp^2 + (\mu_\parallel^2-\mu_\perp^2)(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{K}}})^2}}, \tag{C3}$$

and

$$(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{R}}}) = \frac{\mu_\parallel(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{K}}})}{\sqrt{\mu_\perp^2 + (\mu_\parallel^2-\mu_\perp^2)(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{K}}})^2}}. \tag{C4}$$

Combination of these two equations leads to a direct relationship between $\hat{\underline{\mathbf{K}}}$ and $\hat{\underline{\mathbf{R}}}$ according to

$$\hat{\underline{\mathbf{K}}} = \frac{\mu_\parallel\hat{\underline{\mathbf{R}}} + (\mu_\perp-\mu_\parallel)(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{R}}})\underline{\mathbf{a}}}{\sqrt{\mu_\parallel^2 + (\mu_\perp^2-\mu_\parallel^2)(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{R}}})^2}}; \tag{C5}$$

also the modulus of group velocity can be obtained as a spatial function as

$$c_{SH}(\hat{\underline{\mathbf{R}}}) = \sqrt{\frac{\mu_\perp\mu_\parallel}{\varrho(\mu_\parallel + (\mu_\perp-\mu_\parallel)(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{R}}})^2)}}. \tag{C6}$$

APPENDIX D. GREEN'S TRIADIC FUNCTION

The *SH* part of Green's triadic function Σ_{SH} is defined as[1]

$$\underline{\underline{\Sigma}}_{SH}(\underline{\mathbf{R}},\omega) = \underline{\underline{\mathbf{C}}}:\underline{\nabla}\,\underline{\mathbf{G}}_{SH}(\underline{\mathbf{R}},\omega) \tag{D1}$$

and is symmetrical in the first two indices. Using Eqs. (B1)–(B4) and obeying $R\to\infty$ allows us to perform the $\underline{\nabla}$ operations on $\underline{\mathbf{G}}_{SH}$ according to Eq. (D1). Additionally it is

$$\underline{\nabla}\,\hat{\underline{\mathbf{u}}}(\hat{\underline{\mathbf{R}}}) = -\frac{1}{R[1-(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{R}}})^2]^{1/2}}\left[\frac{(\hat{\underline{\mathbf{R}}}-(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{R}}})\underline{\mathbf{a}})(\underline{\mathbf{a}}\times\hat{\underline{\mathbf{R}}})}{[1-(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{R}}})^2]} + (\underline{\mathbf{a}}\times\underline{\underline{\mathbf{I}}})\right], \tag{D2}$$

where, with $\underline{\mathbf{a}} = (a_x, a_y, a_z)$, the antisymmetric matrix is given by

$$\underline{\mathbf{a}}\times\underline{\underline{\mathbf{I}}} = \begin{pmatrix} 0 & a_z & -a_y \\ -a_z & 0 & a_x \\ a_y & -a_x & 0 \end{pmatrix}. \tag{D3}$$

In the far-field approximation the derivatives according to Eq. (D2) can be neglected, since only terms of the order $(1/R)$ are retained. Thus it is

$$\underline{\underline{\Sigma}}_{SH}^{\mathrm{far}}(\underline{\mathbf{R}},\omega) = j\omega\tau\,\frac{e^{j\omega R/c_{SH}(\hat{\underline{\mathbf{R}}})}}{4\pi R} \times\underline{\underline{\mathbf{C}}}:(\mu_\parallel\hat{\underline{\mathbf{R}}} + (\mu_\perp-\mu_\parallel)\underline{\mathbf{a}})\hat{\underline{\mathbf{u}}}(\hat{\underline{\mathbf{R}}})\hat{\underline{\mathbf{u}}}(\hat{\underline{\mathbf{R}}}), \tag{D4}$$

where

$$\tau = \frac{\sqrt{\varrho/\mu_{\parallel}}}{\mu_{\perp}(\mu_{\parallel}+(\mu_{\perp}-\mu_{\parallel})(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{R}}})^2)}. \tag{D5}$$

The performance of the double contraction in (D4) is straightforward; the final result for the space–time far-field representation of Green's triadic function for cross-plane shear waves in TIA media is

$$\underline{\underline{\Sigma}}_{SH}^{\mathrm{far}}(\underline{\mathbf{R}},t) = \frac{\sqrt{\varrho\,\mu_{\parallel}}}{\mu_{\parallel}+(\mu_{\perp}-\mu_{\parallel})(\underline{\mathbf{a}}\cdot\hat{\underline{\mathbf{R}}})^2}\,(\hat{\underline{\mathbf{u}}}(\hat{\underline{\mathbf{R}}})\hat{\underline{\mathbf{R}}}\hat{\underline{\mathbf{u}}}(\hat{\underline{\mathbf{R}}}) + \hat{\underline{\mathbf{R}}}\hat{\underline{\mathbf{u}}}(\hat{\underline{\mathbf{R}}})\hat{\underline{\mathbf{u}}}(\hat{\underline{\mathbf{R}}}))\,\frac{\delta'(t-R/c_{SH}(\hat{\underline{\mathbf{R}}}))}{4\pi R}. \tag{D6}$$

[1] Y. H. Pao and V. Varatharajulu, ''Huygens' principle, radiation conditions, and integral formulas for the scattering of elastic waves,'' J. Acoust. Soc. Am. **59**, 1361–1371 (1976).

[2] R. G. Payton, *Elastic Wave Propagation in Transversely Isotropic Media* (Martinus Nijhoff, The Hague, 1983).

[3] A. Tverdokhlebov and J. L. Rose, ''On Green's functions for elastic waves in anisotropic media,'' J. Acoust. Soc. Am. **83**, 118–121 (1988).

[4] A. Tverdokhlebov and J. L. Rose, ''On the application domain of the Green's function approximation for mild anisotropic media,'' J. Acoust. Soc. Am. **86**, 1606–1607 (1989).

[5] V. K. Tewary and C. M. Fortunko, ''A computationally efficient representation for propagation of elastic waves in anisotropic solids,'' J. Acoust. Soc. Am. **91**, 1888–1896 (1992).

[6] C. Y. Wang and J. D. Achenbach, ''A new look at 2-D time-domain elastodynamic Green's functions for general anisotropic solids,'' Wave Motion **6**, 389–405 (1992).

[7] W. Ren, ''Spherical wave functions and dyadic Green's functions for homogeneous elastic anisotropic media,'' Phys. Rev. E **47**, 4439–4446 (1993).

[8] J. H. M. T. van der Hijden, *Propagation of Transient Elastic Waves in Stratified Anisotropic Media* (North-Holland, Amsterdam, 1987).

[9] M. Spies, ''Elastic wave propagation in general transversely isotropic media. I: Green's functions and elastodynamic holography,'' J. Acoust. Soc. Am. **96**, 1144–1157 (1994).

[10] K. Helbig, *Foundations of Anisotropy for Exploration Seismics* (Pergamon, New York, 1994).

[11] M. G. Silk, ''The generation of polarized shear waves and their propagation in steels,'' in *Proceedings of the International Symposium on New Methods in NDT and Their Application in Nuclear Engineering*, Saarbrücken, Germany, 17–19 September 1979 (DGZFP, Berlin, 1980), pp. 467–480.

[12] R. B. Thompson, S. S. Lee, and J. F. Smith, ''Absolute measurement of stress in textured plates from angular dependence of *SH* velocity,'' in *Review of Progress in QNDE*, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York, 1984), Vol. 3, pp. 1311–1319.

[13] C. M. Fortunko, ''Ultrasonic evaluation of austenitic stainless steel welds using shear horizontal waves,'' Appl. Phys. Lett. **39**, 699–700 (1981).

[14] G. Hübschen, ''UT with *SH* waves and electromagnetic ultrasonic (EMUS) transducers,'' in *Review of Progress in QNDE*, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York, 1990) Vol. 9, pp. 815–822.

[15] M. Spies and M. Kröning, ''A computationally efficient modeling code for SH wave propagation in austenitic welds using an explicit space–time Green's function,'' in *Review of Progress in QNDE*, edited by D. O. Thompson and D. E. Chimenti (Plenum Press, New York, 1996) Vol. 15, pp. 145–152.

[16] M. Spies, ''Computationally efficient *SH*-wave modeling in transversely isotropic media,'' J. Nondestruct. Eval. (to be submitted).

[17] M. Spies, ''Elastic waves in homogeneous and layered transversely isotropic media: Plane waves and Gaussian wave packets. A general approach,'' J. Acoust. Soc. Am. **95**, 1748–1760 (1994).

[18] M. Spies, ''Elastic waves in transversely isotropic media: Plane waves, Gaussian wave packets, Green's functions, elastic holography,'' Ph.D. thesis, University of Saarbrücken, Germany, 1992 (in German).

[19] This approach has been used by Chen (see Ref. 20) for the case of a uniaxial electromagnetic medium; another possible method has been applied in Ref. 9.

[20] H. C. Chen, *Theory of Electromagnetic Waves—A Coordinate-Free Approach* (McGraw-Hill, New York, 1983).

[21] V. T. Buchwald, ''Elastic waves in anisotropic media,'' Proc. R. Soc. London, Ser. A **253**, 563–580 (1959).

[22] A. Ben-Menahem and R. L. Gibson, ''Directional attenuation of *SH* waves in anisotropic poroelastic inhomogeneous media,'' J. Acoust. Soc. Am. **93**, 3057–3065 (1993).

[23] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products* (Academic, New York, 1980).

[24] F. Oberhettinger, *Fourier Transforms of Distributions and Their Inverses* (Academic, New York, 1973).

Focusing of *N* waves in air by an ellipsoidal reflector

Wayne M. Wright[a] and David T. Blackstock
Applied Research Laboratories, The University of Texas at Austin, P.O. Box 8029, Austin, Texas 78713-8029

(Received 14 June 1996; accepted for publication 31 March 1997)

An experimental investigation of the focusing of intense, airborne pressure pulses by an ellipsoidal reflector is reported. Short-duration *N* waves were generated by weak sparks at the near focus F_1 of four different prolate ellipsoidal surfaces. Reflection then concentrated the rays at the far focus F_2. Measurements were made with a wide-band microphone, primarily along the axis of the reflector. A few transverse measurements in the plane of F_2 were also made. In the axial measurements the reflected (upright *N* wave) and edge-diffracted (inverted *N* wave) components of the signal are distinct at points distant from F_2, approach each other as the focus is neared, and merge at the focus. At the focus the waveform is U-shaped, or cusped. Beyond F_2 a ''droopy'' *N* wave emerges and, where the edge wave can be resolved, it now arrives first and is phase inverted relative to its prefocal version. The more intense reflected wave does not experience phase inversion. Some features of the observed waveforms are consistent with recent calculations by Hamilton, which are based on lossless, linear theory. Other wave characteristics, including pulse amplitude and location of the head and tail shocks, are strongly influenced by nonlinear effects. [S0001-4966(97)04708-5]

PACS numbers: 43.25.Jh, 43.25.Zx, 43.20.Px, 43.80.Sh [MAB]

INTRODUCTION

Focusing by ellipsoidal reflectors, demonstrated by measurements with short-duration pressure pulses in air, is the subject of this paper. The use of pulses instead of cw signals helps identify various components of the wave field. In our case the pulses are *N* waves generated by weak electrical sparks. Airborne *N* waves have also been used to study focusing by a spherical mirror (see, for example, Cornet[1]) and by a paraboloidal reflector (see, for example, Beasley *et al.*[2]). Other surface contours have advantages for particular applications. The ellipsoidal surface provides excellent transfer of wave energy between two points (or foci) of moderate separation. Ellipsoidal reflectors find practical use in lithotripsy,[3–6] an important recent application of acoustical technology to medicine, in which the fluid medium consists of water and body tissue. Indeed, questions raised about lithotriper sound fields motivated the present study.

Extracorporeal shock wave lithotripters employ a source of intense acoustic waves, located outside the human body, to destroy kidney stones and gallstones *in situ* without (it is hoped) causing appreciable damage to living tissue. In one widely used commercial lithotripter, from the German firm Dornier,[5,6] energetic sparks in water produce (nearly unipolar) positive pressure pulses at the interior focus F_1 of a hemi-ellipsoidal reflector, and the patient's body is positioned so that the stone is located at the second focus F_2. Irradiation by successive focused pulses causes erosion and eventual breakup of the stone into fragments small enough to pass out of the body naturally through the urinary tract.

Theoretical modeling of the ellipsoidal sound field is challenging. Three factors—focusing, nonlinear distortion, and diffraction—interact in a complicated way to produce the characteristic lithotripter pulse shape at the focus.[5] An analytical solution has been obtained by Hamilton[7] for the small-signal problem (nonlinear effects and losses neglected). He used the Helmholtz–Kirchhoff integral to predict waveforms along the axis of an ellipsoidal reflector. Christopher[6] has developed a computational model for ellipsoidal focusing of finite-amplitude waves (effects of absorption and dispersion also included). His predictions, specifically for a Dornier HM3 lithotripter, agree well with the corresponding measurements of Coleman *et al.*[5] When absorption and nonlinearity are removed from the algorithm, he obtains excellent agreement with Hamilton's on-axis analytical waveforms. When nonlinearity is included, his computed waveforms near F_2 differ substantially in both shape and amplitude from the small-signal waveforms.

Most experimental studies of ellipsoidal focusing have been done in water or tissue. By performing our study in air, we avoided the complication of cavitation and the bubble pulse. Moreover, the nearly symmetric *N* wave produced by a spark in air is in some ways better suited for interpretation and analysis than is the waterborne spark-produced pulse. That pulse is a very sharp positive spike followed by a long shallow negative tail.

Most of the measurements reported here are along the axis, but a few are transverse, in the plane of F_2. The effects of focusing, diffraction, and nonlinearity are clearly evident in the waveforms.

I. ELLIPSOIDAL REFLECTION

Figure 1 indicates the pertinent definitions and geometry for the prolate ellipsoidal reflectors used in this study. The origin of the r,z cylindrical coordinate system is at the second focus F_2; the center of the ellipsoid is at $z = -z_0$. The

[a]Permanent address: Department of Physics, Kalamazoo College, Kalamazoo, MI 49006, Electronic mail: wwright@kzoo.edu

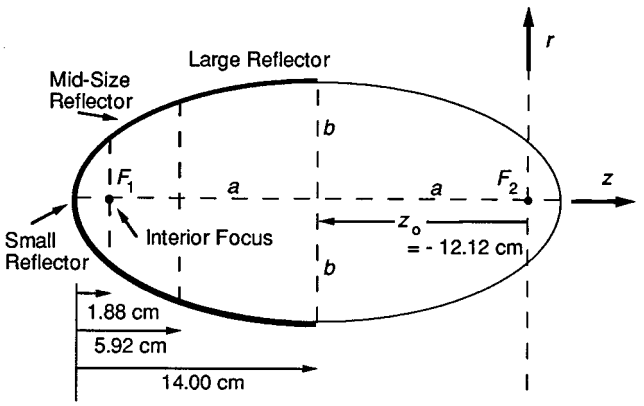


FIG. 1. Geometry of ellipsoid, showing interior (F_1) and exterior (F_2) foci and three reflectors with differing depths but common axes ($2a=28$ cm, $2b=14$ cm) and eccentricity $\epsilon = 0.866$.

equation of the surface is $r^2/b^2+(z+z_0)^2/a^2=1$, the lengths of the major and minor axes are $2a$ and $2b$, respectively, and $\epsilon = \sqrt{1-(b/a)^2}$ is the eccentricity. Regardless of the launch direction, sound originating at focus F_1, arrives after a single reflection at the second focus F_2 after having traveled a path of length $2a$. If the waves are small signal, therefore, the travel time from F_1 to F_2 is independent of the particular path followed.

Since the reflectors used in this experiment are only portions of an ellipsoidal surface, for example, the section to the left of one of the vertical dashed lines in Fig. 1, the fraction of energy actually reaching F_2 from F_1 is proportional to the solid angle subtended by the surface at F_1. Diffraction by the edge of the reflector is expected to produced an edge wave. The spatial relation of the edge wave to the reflected wave is indicated in Fig. 2, which shows representative reflected wavefronts *B* and edge wavefronts *C*. It is seen that if the receiver is on-axis in front of the focus ($z<0$), the reflected wave leads the edge wave. The reverse is true if the receiver is beyond the focus ($z>0$). Only at focus F_2 do both waves arrive simultaneously. From linear theory one expects (1) the edge wave on axis to be phase inverted relative to the reflected wave, and (2) the reflected wave to suffer phase inversion as it passes through the focus. Thus if the prefocal reflected signal is a conventional *N* wave, the edge wave there is an upside down *N*; beyond the focus the polarities of the two signals reverse. If finite-amplitude effects

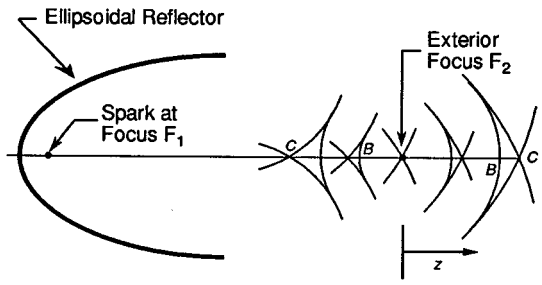


FIG. 2. Schematic representation of reflected (*B*) and edge-diffracted (*C*) wavefronts from a hemi-ellipsoidal reflector.

are important, nonlinear distortion of the reflected wave should be particularly consequential in the vicinity of the focal point. Indeed, finite-amplitude effects may prevent, or at least modify, phase inversion of the reflected wave.[8]

Another expectation has to do with the magnitude of the edge wave relative to that of the reflected wave. The rays striking the rim of the reflector are weaker than the paraxial reflected rays, since on the ellipsoidal surface the ray density is least at the rim, and thus the edge wave is relatively weak. The weakness of the edge wave is expected to be more pronounced for deep reflectors than for shallow ones.

II. EXPERIMENT

Four ellipsoidal reflectors were machined from aluminum blocks. Three—denoted Large, Mid-Size, and Small in Fig. 1—all have the same major axis $2a=28$ cm, minor axis $2b=14$ cm, and eccentricity $\epsilon=0.866$. The depths are given in Fig. 1. The fourth reflector has the same separation between foci but eccentricity $\epsilon=0.753$ (other measurements: $2a=32.2$ cm, $2b=21.2$ cm, depth behind $F_1=4.0$ cm). The large reflector is the deepest one. It is a hemi-ellipsoid; the separation between its aperture plane and either focus is 12.12 cm. Although the fourth reflector has the same aperture size and distance between foci as the large reflector, it is much more shallow; its aperture is in the plane of F_1. The reason for making four different reflectors for the experiment was that we wished to test some of the expectations described in the previous paragraph, particularly the properties of the edge wave.

The acoustic source for each reflector was the same, a weak spark produced at F_1 between pointed tungsten electrodes having nominal 3-mm separation. The parallel combination of a high-voltage power supply (Glassman series EH), a 1.0-nF mica capacitor, and the spark gap functions as a simple relaxation oscillator. Firing voltage was about 5 kV. Adjustment of the spark repetition rate was possible, through control of the supply voltage, up to several discharges per second.[9]

The microphone was of the wide-band electrostatic type.[10] The diaphragm was 1/4-mil (6-μm) aluminized Mylar, stretched over a 16-mm-diam cartridge which contained the 1.6-mm-diam smooth metal back electrode, embedded in a plastic insulator. To avoid diffraction by the microphone itself, we fit the cartridge flush in a 127-mm-diam plane baffle.[9] The microphone was attached directly to a preamplifier, which provided a high-impedance load as well as a 135-V dc bias voltage for the microphone. Individual waveforms were stored on a Tektronix RTD 710A Digitizer. The baffled microphone had a free-field response of approximately 0.05 mV/Pa from low frequencies to above 0.5 MHz.

A sturdy optical bench was used for mounting the spark electrodes, ellipsoidal reflector, and microphone. Positioning was done manually. Alignment was maintained with a HeNe laser beam.

Careful measurement technique was required. Near the exterior focus F_2, the measured waveforms were found to be especially sensitive to shifts in spark location. Successive sparks do not follow precisely the same path between pointed electrodes, and a spark 3 mm in length is not truly a

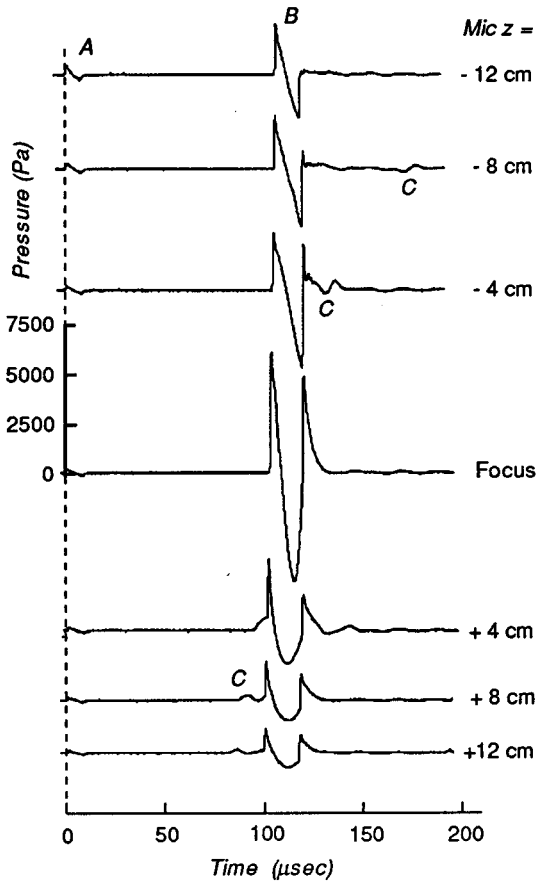


FIG. 3. Waveforms observed along the axis of the large hemi-ellipsoidal reflector, with aperture diameter 14 cm, depth 14.0 cm, and eccentricity 0.866. The time origin for each position corresponds to arrival of the direct N wave from the spark (A). Reflected (B) and edge-diffracted (C) signals are indicated.

point acoustic source. Even when the experimental apparatus was aligned with great care, variations in amplitude at the focus of up to 20% were not unusual.

III. EXPERIMENTAL RESULTS

Figure 3 shows axial waveforms for the large reflector. The pressure scale is the same for all traces. A time reference is provided by the direct N wave from the spark, labeled A. For the top waveform ($z=-12$ cm), the microphone is midway between the foci, in the aperture plane of the reflector. One notes that the reflected signal B is initially an ordinary N wave (similar to A), while the edge wave, labeled C, is inverted. As the microphone approaches the focus F_2, the separation between the edge and reflected waves decreases, but the two signals are still fully resolved 4 cm in front of the focus. For both pulses, the amplitude increases slightly as F_2 is approached.

Reflected and diffracted waves merge to produce a cusped, or U-shaped, wave at the focus. The pressure amplitude there is greater than 5000 Pa=0.05 atm. At larger axial distances the waveform takes on a ''droopy'' N wave appearance, a kind of blend of U and N shapes. Two shocks are evident, but the curve connecting them is not linear. The second shock overshoots into the positive pressure region and has a tail. Beyond F_2 the diffracted pulse C arrives ahead of the reflected one, as promised by Fig. 2. Moreover, it appears inverted relative to its prefocal shape, beginning with a peak rather than a trough. Also to be noticed in the sequence of waveforms is the gradual increase in separation of the head and tail shocks; this is a finite-amplitude effect.

Similar sequences are shown for the mid-size reflector in Fig. 4(a) and for the small reflector in Fig. 4(b). Source conditions were not quite the same as those used for the large reflector. The sparks were slightly more intense. Because of the relative weakness of the small reflector field, the signals in Fig. 4(b) have been magnified by a factor of 2 relative to those in Fig. 4(a). Although no pressure scale is given (it is not known to the same accuracy as that in Fig. 3), the trend of the waveforms is clear. One sees that the results differ considerably from those obtained for the large reflector. The edge wave C is relatively stronger, but because of the smaller aperture radius, signals C and B have smaller separation and C merges with B more quickly. These results were expected. Failure of an inverted edge wave to emerge in front of the reflected wave for the postfocal region is noteworthy. The amplitude increase of the reflected wave as it approaches F_2 is not as appreciable as in Fig. 3. Nor is the wave at the focus as far along in the transition from N to U shape.

Measurements with the fourth reflector (wide but shallow) are presented in Fig. 5. Axial measurements are shown in Fig. 5(a), transverse measurements in the plane of F_2 in Fig. 5(b). In the axial waveforms the pressure amplitude at the focus is comparable to that measured for the large reflector, while the overall waveform shapes beyond F_2 more closely resemble those for the mid-size reflector. Especially noticeable in Fig. 5(a) are (1) a pronounced increase in shock separation at $z=-4$ cm, where the negative initial portion of the edge wave has merged with the negative second half of the reflected wave, and (2) the gradual lengthening of the wave beyond the focus. The transverse waveform sequence in Fig. 5(b), displayed as left and right pairs, shows that the field was symmetric and the apparatus well-aligned. Since the edge wave is strong (coherent) only on the axis, it does not appear in the Fig. 5(b) traces. The approximate diameter of the focal region, within which the pressure amplitude decreases from the axial value by 6 dB, is 1.6 cm.

IV. COMPARISON WITH LINEAR THEORY

Hamilton[7,11] has evaluated his linear solution to predict the axial pressure waveforms for N waves incident upon a number of ellipsoidal reflectors. For each case he calculates three distinct contributions: a *center wave*, an *edge wave*, and a *wake*. The wake is controlled by a term in the analytical solution that results from the nonuniform pressure amplitude at the mirror surface.

Hamilton's linear analysis (as presented in Ref. 7) and our experiments in air show qualitative agreement as follows:

(1) When the axial observation point is in front of the focus F_2, the edge wave is distinct and inverted and lags behind the center wave. As F_2 is approached the lag decreases. The relative amplitudes of reflected and edge waves

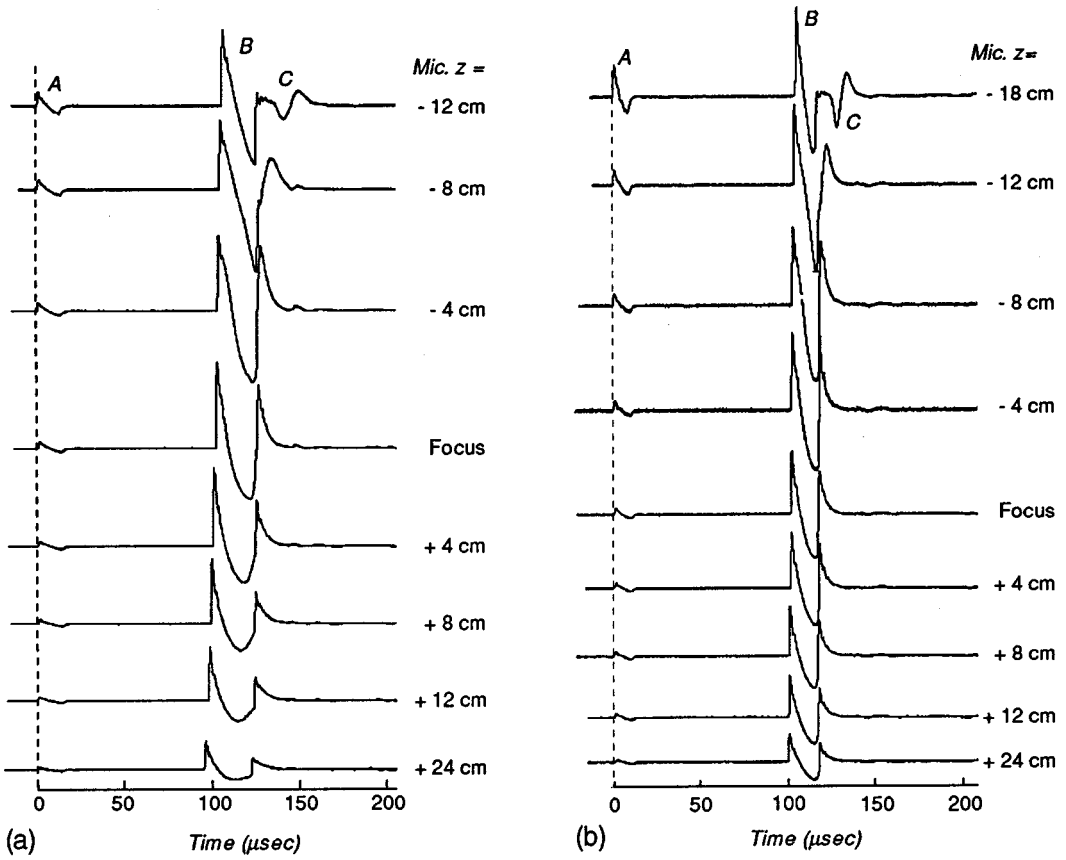


FIG. 4. Waveforms observed along the major axis of (a) the mid-size reflector (depth 5.92 cm) and (b) the small reflector (depth 1.88 cm; with $2x$ gain).

that we observed are in reasonable agreement with predictions. The two waves merge in the focal region, and at greater distances the edge wave leads and, for those experimental situations in which it is resolved, is an upright N. The failure of the reflected wave in our experiments to invert may be attributed to finite-amplitude effects.

(2) For the situations of Figs. 4 and 5, even before the reflected wave reaches focus F_2 the expansion connecting the head and tail shocks ceases be linear and curves upward. The tail shock overshoots into positive pressure and has its own tail. These features are associated with the wake. See, e.g., Fig. 8(h) of Ref. 7.

(3) At the exterior focus the waveform tends to be cusped or U-shaped. The observed shape is consistent with the derivative of the source function that is predicted by Eq. (30) of Ref. 7.

(4) Between the two foci the edge wave for a particular ellipsoidal surface becomes more significant, relative to the reflected wave, as the reflector depth is reduced. This effect corresponds to the fact that rims of shallow reflectors are in regions of higher ray density than rims of deep reflectors.

Although the assumption of linear superposition of wave contributions is not truly valid for the finite-amplitude signals in our experiments, it can still lead to helpful benchmarks. Consider, for example, the amount of radiation that would be incident upon each of the reflectors if the spark were an ideal point source at F_1. The initial solid angle within which the incident wave is specularly reflected is reduced, relative to that for our large reflector, by factors of 0.85 and 0.54 for the mid-size and small reflectors, respectively. The pressure amplitudes measured at F_2 for these two reflectors are about 40% of the values based on these factors. The discrepancy can be attributed largely to nonlinear propagation effects, which cause attenuation of the N wave.

Much evidence of finite-amplitude effects appears in the measured waveforms. The head and tail shocks of the reflected wave become increasingly separated with propagation distance. Moreover, notice the difference between the duration of the much weaker direct wave A and the reflected wave B, even far before the focus. Attention has already been called to the failure of the reflected wave to invert when it passes through the focus. In the postfocal region, steepening of the compression segments of the waveform, and corresponding easing of the expansion segments, contribute to the gradual change from U shape in the focal region to the ''droopy'' N shape beyond the focus.

V. CONCLUSIONS

Short-duration acoustic transients (N waves), produced by weak sparks at the interior focus F_1, have been reflected from four different ellipsoidal mirrors. As the microphone is moved along the mirror axis toward the exterior focus F_2, the initial N wave distorts somewhat toward a U-shaped, or cusped, wave. Shocks are present in all of the observed waveforms.

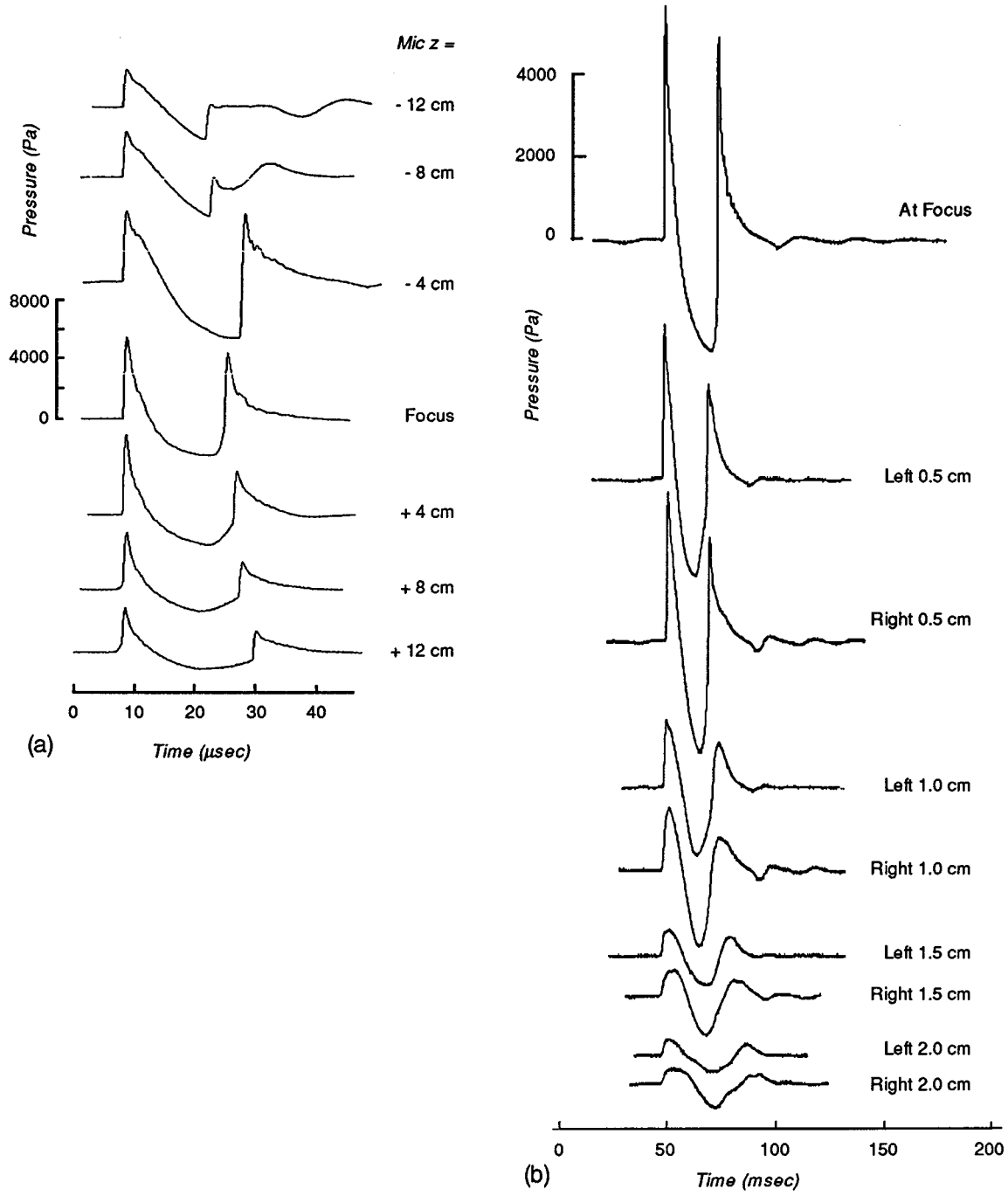


FIG. 5. Waveforms observed (a) along the axis and (b) along a diameter through the exterior focus F_2 of the fourth ellipsoidal reflector. Aperture diameter 14 cm, depth 4.0 cm, and eccentricity 0.753.

A distinct diffracted wave, which originates at the edge of the reflector and propagates with leading rarefaction, is observed in front of the focus for each of the four reflectors. Only for the largest reflector did we observe the edge wave in the postfocal region. In that case the edge wave begins, as predicted, with a condensation.

Numerical calculations by Hamilton, based on a lossless, linear propagation model, predict waveforms that compare favorably in several respects with our observations. Particularly noteworthy is the observed distortion of the N wave that is attributable to the wake term in the analysis.

The most interesting finite-amplitude effects are the lengthening of the signal between shocks and the failure of the reflected wave to become an inverted N beyond the focus F_2. The computational work of Christopher supports our finding that neglect of nonlinear propagation effects can lead to a substantial overestimate for the amplitude in the focal region.

ACKNOWLEDGMENTS

We gratefully acknowledge the assistance of Fred Cotaras, Matthew Jones, Stephen Cheng, and Pingwah Li in the

design and construction of the ellipsoidal reflectors. The extra small-signal calculations made by Mark Hamilton and the many discussions he provided were particularly helpful. This work received partial support from the Office of Naval Research and from the IR&D Program of the Applied Research Laboratories, The University of Texas at Austin.

[1] E. P. Cornet, ''Focusing of an *N* wave by a spherical mirror,'' Tech. Rep. ARL-TR-72-40, Applied Research Laboratories, The University of Texas at Austin, September 1972.
[2] W. D. Beasley, J. D. Brooks, and R. L. Barger, ''A laboratory investigation of *N* wave focusing,'' NASA Technical Note TND-5306, Langley Research Center, Hampton, VA, 1969.
[3] M. Müller, ''Experimental investigation into the focusing of spherical shock waves in water using deep ellipsoidal reflectors'' (in German), Acustica **66**, 258–266 (1988).
[4] A. J. Coleman and J. E. Saunders, ''A survey of the acoustic output of commercial extracorporeal shock wave lithotripters,'' Ultrasound Med. Biol. (UK) **15**, 213–227 (1989).
[5] A. J. Coleman *et al.*, ''Pressure waveforms generated by a Dornier extracorporeal shock-wave lithotripter,'' Ultrasound Med. Biol. (UK) **13**, 651–657 (1987).
[6] T. Christopher, ''Modeling the Dornier HM3 lithotripter,'' J. Acoust. Soc. Am. **96**, 3088–3095 (1994).
[7] M. F. Hamilton, ''Transient axial solution for the reflection of a spherical wave from a concave ellipsoidal mirror,'' J. Acoust. Soc. Am. **93**, 1256–1266 (1993).
[8] B. Sturtevant and V. A. Kulkarny, ''The focusing of weak shock waves,'' J. Fluid Mech. **73**, 651–671 (1976).
[9] W. M. Wright, ''Propagation in air of *N* waves produced by sparks,'' J. Acoust. Soc. Am. **73**, 1948–1955 (1983).
[10] W. M. Wright, ''High frequency electrostatic transducers for use in gases,'' Tech. Memo. No. 47, Acoustics Research Laboratory, Harvard University, April 1962.
[11] M. F. Hamilton (personal communication).

Dissipative effects on Bjerknes forces between two bubbles

Alexander A. Doinikov
Institute of Nuclear Problems, Bobruiskaya 11, 220050 Minsk, Byelorussia

(Received 26 June 1996; revised 8 March 1997; accepted 25 March 1997)

The time-averaged forces exerted by a sound field between two gas bubbles, normally called Bjerknes forces, are examined. The surrounding medium is assumed to be an incompressible viscous liquid. It is shown that the acoustic streaming, which develops around the bubbles due to viscosity, adds one more term to the Bjerknes force that is dominant when the viscous wavelength is large compared with the bubble radii. © *1997 Acoustical Society of America.*
[S0001-4966(97)05007-8]

PACS numbers: 43.25.Yw [MAB]

INTRODUCTION

Bjerknes forces arising between two bubbles in a sound field and causing them to attract or repel each other are well-known in acoustics. They are named for C. A. Bjerknes and his son V. F. K. Bjerknes, who were the first to report on this effect and to derive an analytical expression for the forces.[1] The later literature on this topic is reviewed in sufficient detail in several articles[2–5] and therefore only works directly relevant to the subject matter of the present paper will be mentioned below.

The purpose of this paper is to study dissipative effects on the Bjerknes forces between two radially pulsating bubbles. When deriving an expression for the forces, C. A. Bjerknes is known to have neglected the dissipative effects. These appeared to be first taken into account by Zabolotskaya[6] using the well-known phenomenological procedure: She entered damping into equations for the instantaneous radii of the bubbles that were initially obtained for an *ideal* liquid. This approach, however, takes no account of the acoustic streaming which develops around the bubbles due to the viscosity of the surrounding liquid. As is shown in this paper, the acoustic streaming gives rise to one more term in the Bjerknes force that plays an important role when the viscous effects become strong.

I. THEORY

A. Problem formulation

Let two gas bubbles with mean radii R_1 and R_2, separated by a distance l, be immersed in a liquid irradiated by a sound wave field. We assume the wavelength of the sound field to be much larger than R_1, R_2, and l. It is also assumed that thermal effects in the liquid are negligible, which is the case when the thermal penetration depth is small compared with R_1 and R_2. These assumptions allow us to neglect the compressibility of the liquid and represent equations of the liquid motion[7] and the mean force on each bubble[8] as follows:

$$\nabla\cdot\mathbf{v}=0, \tag{1}$$

$$\rho\,\frac{\partial\mathbf{v}}{\partial t}+\rho(\mathbf{v}\cdot\nabla)\mathbf{v}=-\nabla p+\eta\Delta\mathbf{v}, \tag{2}$$

$$\mathbf{F}_j=\int_{S_j}\langle 2\eta(\mathbf{n}_j\cdot\nabla)\mathbf{v}^{(2)}+\eta\mathbf{n}_j\times(\nabla\times\mathbf{v}^{(2)})-\mathbf{n}_j p^{(2)} -\rho\mathbf{v}^{(1)}(\mathbf{v}^{(1)}\cdot\mathbf{n}_j)\rangle dS_j, \tag{3}$$

where $\mathbf{v}$ is the liquid velocity, ρ is the liquid density, p is the liquid pressure, η is the shear viscosity of the liquid, $\mathbf{F}_j$ is the mean force on the jth bubble ($j=1,2$), S_j is the equilibrium surface of the jth bubble, $\mathbf{n}_j$ is the unit outward normal to S_j, the superscripts (1) and (2) in Eq. (3) denoted quantities of first and second order in the amplitude of the incident field, and $\langle\ \rangle$ means an average over the incident field cycle. As Eq. (3) shows, calculation of the force requires $\mathbf{v}^{(1)}$, $\langle\mathbf{v}^{(2)}\rangle$, and $\langle p^{(2)}\rangle$. The first-order velocity $\mathbf{v}^{(1)}$ is determined by the linearized equations (1) and (2)

$$\nabla\cdot\mathbf{v}^{(1)}=0, \tag{4}$$

$$\rho\,\frac{\partial\mathbf{v}^{(1)}}{\partial t}+\nabla p^{(1)}-\eta\Delta\mathbf{v}^{(1)}=0. \tag{5}$$

Assuming that $l\gg R_1,R_2$, we can treat pulsations of the bubbles as purely radial and accordingly represent $\mathbf{v}^{(1)}$ as follows:

$$\mathbf{v}^{(1)}=\nabla\varphi, \tag{6}$$

where $\varphi=\varphi_I+\varphi_1+\varphi_2$ in which $\varphi_I=A\exp(-i\omega t)$ is the incident potential,

$$\varphi_j=A\exp(-i\omega t)a_jR_j/r_j \tag{7}$$

is the scattered potential of the jth bubble, r_j is the radial coordinate of the spherical coordinate system with the origin at the equilibrium center of the jth bubble (see Fig. 1), and a_1 and a_2 are dimensionless constants to be determined by the boundary conditions at the surfaces of the bubbles. We will find expressions for these constants later; for the moment we assume them to be known and proceed to calculate $\langle\mathbf{v}^{(2)}\rangle$ and $\langle p^{(2)}\rangle$.

B. Equations of acoustic streaming

By averaging Eqs. (1) and (2) over time and keeping up to the second order, we obtain equations for $\langle\mathbf{v}^{(2)}\rangle$ and $\langle p^{(2)}\rangle$ which are normally called equations of acoustic streaming

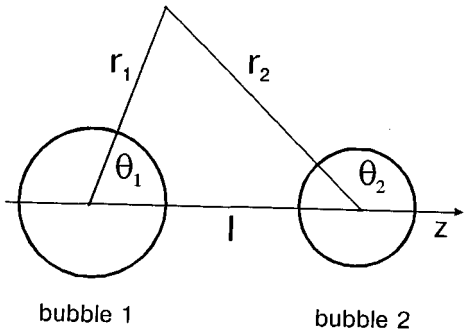


FIG. 1. Illustration of the coordinate systems used.

$$\nabla \cdot \langle \mathbf{v}^{(2)} \rangle = 0, \tag{8}$$

$$\eta \Delta \langle \mathbf{v}^{(2)} \rangle - \nabla \langle p^{(2)} \rangle = \rho \langle (\mathbf{v}^{(1)} \cdot \nabla) \mathbf{v}^{(1)} \rangle. \tag{9}$$

We seek $\langle \mathbf{v}^{(2)} \rangle$ as

$$\langle \mathbf{v}^{(2)} \rangle = \nabla \mathbf{\Phi} + \nabla \times \mathbf{\Psi}. \tag{10}$$

Substitution of Eqs. (6) and (10) into Eqs. (8) and (9) yields

$$\Delta \mathbf{\Phi} = 0, \tag{11}$$

$$\Delta^2 \mathbf{\Psi} = 0, \tag{12}$$

$$\nabla \langle p^{(2)} \rangle = \eta \Delta (\nabla \times \mathbf{\Psi}) - \tfrac{1}{2} \rho \nabla \langle (\nabla \varphi)^2 \rangle. \tag{13}$$

We shall calculate the force on bubble 1. (The force on bubble 2 can then be found replacing quantities concerning bubble 1 by similar quantities concerning bubble 2 and *vice versa*.) In view of axial symmetry, we seek $\mathbf{\Phi}$ and $\mathbf{\Psi}$ as

$$\mathbf{\Phi} = \alpha (R_1 / r_1)^2 \cos \theta_1, \tag{14}$$

$$\mathbf{\Psi} = \beta \mathbf{e}_{\epsilon_1} \sin \theta_1, \tag{15}$$

where $\mathbf{e}_{\epsilon_1}$ is the unit vector of the spherical coordinate system related to bubble 1, and α and β are constants. Equations (14) and (15), as written here, assume effects of the acoustic streaming of bubble 2 on the force exerted on bubble 1 to be negligible. It is easy to verify that Eqs. (14) and (15) meet Eqs. (11) and (12) as well as the condition $\langle \mathbf{v}^{(2)} \rangle \to 0$ as $r_1 \to \infty$. To find α and β, we use the boundary condition for the liquid velocity at the surface of bubble 1. Note that the condition should be applied to the *perturbed* surface of bubble 1, whereupon it should be taken with an accuracy up to the second order and averaged over time. To get this result, we represent the perturbed surface of bubble 1 as $r_1 = R_1 + x_1(t)$ and expand the liquid velocity $\mathbf{v}$ at this surface in a Taylor series

$$\begin{aligned} \mathbf{v}|_{r_1 = R_1 + x_1} &\approx \mathbf{v}|_{r_1 = R_1} + x_1 \left. \frac{\partial \mathbf{v}}{\partial r_1} \right|_{r_1 = R_1} \\ &\approx \mathbf{v}^{(1)}|_{r_1 = R_1} + \mathbf{v}^{(2)}|_{r_1 = R_1} + x_1^{(1)} \left. \frac{\partial \mathbf{v}^{(1)}}{\partial r_1} \right|_{r_1 = R_1}, \end{aligned} \tag{16}$$

where $x_1^{(1)}$ is the first-order change in the radius of bubble 1, subject to the following relation: $\partial x_1^{(1)} / \partial t = v_{r_1}^{(1)}$ at $r_1 = R_1$. When averaged over time, Eq. (16) gives

$$\langle \mathbf{v}|_{r_1 = R_1 + x_1} \rangle = \langle \mathbf{v}^{(2)} \rangle|_{r_1 = R_1} + \left\langle x_1^{(1)} \left. \frac{\partial \mathbf{v}^{(1)}}{\partial r_1} \right|_{r_1 = R_1} \right\rangle. \tag{17}$$

We should also take into consideration that the total steady velocity field in addition to the velocity considered above involves a component due to slow translational motions of the bubbles caused by the Bjerknes forces. Calling this component $\mathbf{U}$ and adding to Eq. (17), the desired boundary condition is given by

$$\mathbf{U} + \langle \mathbf{v}^{(2)} \rangle + \left\langle x_1^{(1)} \frac{\partial \mathbf{v}^{(1)}}{\partial r_1} \right\rangle = \mathbf{u}_1 \quad \text{at } r_1 = R_1, \tag{18}$$

where $\mathbf{u}_1$ is the velocity of the slow translational motion of bubble 1. By its very definition, $\mathbf{U}$ should be subject to the condition: $\mathbf{U} = \mathbf{u}_1$ at $r_1 = R_1$ (as well as $\mathbf{U} = \mathbf{u}_2$ at $r_2 = R_2$) and hence Eq. (18) becomes

$$\langle \mathbf{v}^{(2)} \rangle = \frac{1}{\omega^2} \left\langle \frac{\partial v_{r_1}^{(1)}}{\partial t} \frac{\partial \mathbf{v}^{(1)}}{\partial r_1} \right\rangle \quad \text{at } r_1 = R_1. \tag{19}$$

This equation means that when calculating the Bjerknes forces, we treat the bubbles as if they were immobile in the second approximation. Substituting Eqs. (6), (7), (10), (14), and (15) into Eq. (19) and retaining only the leading terms in the small parameters R_1/l and R_2/l, one finds $\beta = -\alpha$ and

$$\alpha = \frac{|A|^2}{4 \omega R_1^2} \frac{R_1 R_2}{l^2} \operatorname{Im}(a_1 a_2^*), \tag{20}$$

where Im denotes ''the imaginary part of,'' and the asterisk indicates the complex conjugate. Substituting Eq. (15) into Eq. (13), we obtain

$$\langle p^{(2)} \rangle = 2 \eta \beta \frac{\cos \theta_1}{r_1^2} - \frac{1}{2} \rho \langle (\nabla \varphi)^2 \rangle. \tag{21}$$

Here, we have also used that $\nabla \times (\mathbf{e}_{\epsilon_1} r_1^{-2} \sin \theta_1) = -\nabla (r_1^{-2} \cos \theta_1)$ and that $\langle p^{(2)} \rangle \to 0$ as $r_1 \to \infty$.

C. Calculation of the Bjerknes force

Substituting Eqs. (6), (7), (10), (14), (15), (20), and (21) into Eq. (3) and retaining only the leading term, one finds

$$F_1 = 2 \pi \rho |A|^2 \frac{R_1 R_2}{l^2} \left[\operatorname{Re}(a_1 a_2^*) + \frac{1}{2} \left(\frac{\delta_v}{R_1} \right)^2 \operatorname{Im}(a_1 a_2^*) \right], \tag{22}$$

where F_1 is the magnitude of $\mathbf{F}_1$, including sign, so that $F_1 > 0$ corresponds to attraction of bubble 1 to bubble 2, and $F_1 < 0$ to repulsion, $\delta_v = \sqrt{2 \eta / \rho \omega}$ is the viscous penetration depth, and Re denotes ''the real part of.'' It is seen from Eq. (22) that it will suffice to calculate a_1 and a_2 in the zero approximation in R_1/l and R_2/l, i.e., as if the other bubble were absent. This result is well-known (see, for example, Ref. 9) and for the case considered here is represented as follows:

$$a_j = (\omega_j^2 / \omega^2 - 1 - i \delta_j)^{-1}, \tag{23}$$

where

$$\omega_j = \left(\frac{3\kappa_j P_{gj}}{\rho R_j^2} - \frac{2\sigma}{\rho R_j^3} \right)^{1/2} \tag{24}$$

is the resonance frequency of the jth bubble,

$$\delta_j = \frac{\omega R_j}{c} + 2\left(\frac{\delta_v}{R_j} \right)^2 + \left(\frac{\omega_j}{\omega} \right)^2 d_{tj} \tag{25}$$

is the total dimensionless damping constant, $P_{gj} = P_0 + 2\sigma/R_j$ is the equilibrium gas pressure inside the jth bubble, $P_0 = P_A + \rho g Z$ is the equilibrium pressure of the host liquid, P_A is the atmospheric pressure, g is the gravity, Z is the depth of submergence of the bubbles, and κ_j and d_{tj} are given by

$$\kappa_j = \gamma \left[(1 + d_{tj})^2 \left(1 + \frac{3(\gamma - 1)(\sinh X_j - \sin X_j)}{X_j(\cosh X_j - \cos X_j)} \right) \right]^{-1}, \tag{26}$$

$$d_{tj} = \frac{3(\gamma - 1)[X_j(\sinh X_j + \sin X_j) - 2(\cosh X_j - \cos X_j)]}{X_j^2(\cosh X_j - \cos X_j) + 3(\gamma - 1)X_j(\sinh X_j - \sin X_j)}, \tag{27}$$

where γ is the specific heat ratio of the gas, $X_j = R_j\sqrt{2\omega/D_j}$, $D_j = K/\rho_j c_{pg}$ is the thermal diffusivity of the gas inside the jth bubble, K is the thermal gas conductivity, c_{pg} is the specific gas heat at constant pressure, $\rho_j = \rho_A P_{gj}/P_A$ is the density of the gas inside the jth bubble, and ρ_A is the gas density at the atmospheric pressure P_A.

Upon substitution of Eq. (23) into Eq. (22), we finally obtain

$$F_1 = \frac{2\pi\rho |A|^2 R_1 R_2 G_1}{l^2[(1 - \omega_1^2/\omega^2)^2 + \delta_1^2][(1 - \omega_2^2/\omega^2)^2 + \delta_2^2]}, \tag{28}$$

where

$$G_1 = \Omega + \delta_1\delta_2 + \mathbf{\Delta}_1, \tag{29}$$

$$\Omega = (1 - \omega_1^2/\omega^2)(1 - \omega_2^2/\omega^2), \tag{30}$$

$$\mathbf{\Delta}_1 = \frac{1}{2}\left(\frac{\delta_v}{R_1} \right)^2 \left[\delta_1 \left(\frac{\omega_2^2}{\omega^2} - 1 \right) - \delta_2 \left(\frac{\omega_1^2}{\omega^2} - 1 \right) \right]. \tag{31}$$

Let us compare Eq. (28) with the result of Zabolotskaya.[6] Recall that she entered damping into equations for the instantaneous radii of the bubbles, which were initially derived for an ideal liquid. The comparison shows that in this approach she has lost the term $\mathbf{\Delta}_1$ due to acoustic streaming. This term has two peculiarities: First, it can change sign and in consequence either increase or decrease the net contribution from the dissipative effects, and second, if we find the force on bubble 2, F_2 [for which purpose we should change the subscript 1 in Eqs. (28)–(31) to 2 and vice versa and take Eq. (28) with negative sign], we will see that $F_1 \neq -F_2$ as $\mathbf{\Delta}_1 \neq \mathbf{\Delta}_2$. This result may be explained as follows: In an ideal liquid, the system of two interacting bubbles is conservative, and mechanisms of the occurrence of the Bjerknes forces on each bubble are symmetrical. The Bjerknes force on bubble 1 is caused by interaction of its linear pulsations, induced by the incident field, with the linear scattered field from bubble 2. The same process occurs for bubble 2. Therefore the forces on the bubbles are equal and opposite. In a viscous liquid, both the conservatism and the symmetry of the system are broken. Clearly the acoustic streaming generated by bubble 1 and resulting in an additional force on that bubble is different from that generated by bubble 2 if the bubbles are of unequal size. Therefore the total forces on the bubbles are not equal and opposite to each other. A similar result occurs for a compressible liquid[10] where the conservatism of the system and its symmetry with respect to the incident field are broken by the finite sound speed.

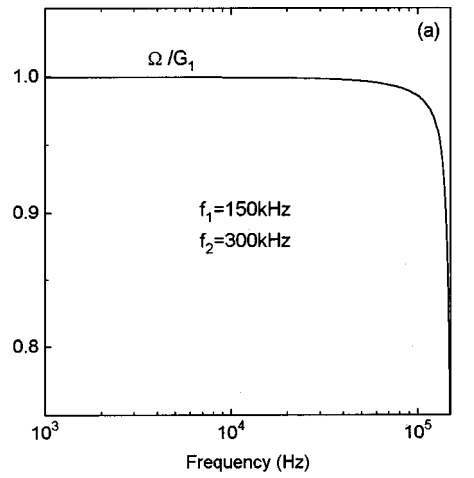
II. NUMERICAL EXAMPLES

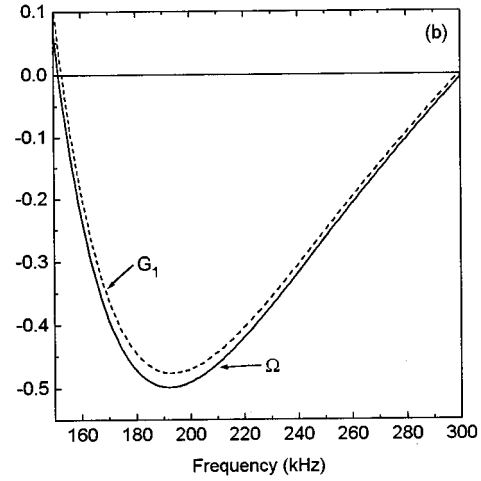
Typically the contribution to the Bjerknes forces from the dissipative effects is insignificant. There are, however, cases where these effects do govern the Bjerknes forces. First, this occurs when ω_1 or ω_2 or both are close to ω, a trivial case. Second, this happens when parameters of the system are such that the viscous wavelength is much larger than the radii of the bubbles. To illustrate both these cases, some numerical calculations have been made.

Figure 2 refers to two air bubbles in water. The corresponding physical parameters are $\rho = 1000$ kg/m^3, $\eta = 10^{-3}$ Pa s, $c = 1500$ m/s, $\sigma = 0.0727$ N/m, $K = 0.034$ J/(sm K), $c_{pg} = 1009$ J/(kg K), $\gamma = 1.4$, $\rho_A = 1.2$ kg/m^3, $P_A = 1.013 \times 10^5$ Pa, and $g = 9.8$ m/s^2. The mean radii of the bubbles were taken to be $R_1 = 20$ μm and $R_2 = 10$ μm, and the depth of their submergence $Z = 0.1$ m. Figure 2(a) exhibits the ratio of the ''ideal'' term of Eq. (29), Ω, to the sum of all the terms, G_1, for frequencies below the resonance frequency f_1 of the larger bubble, Fig. 2(c) shows this ratio for frequencies above the resonance frequency f_2 of the smaller bubble, and Fig. 2(b) presents curves for Ω and G_1 for the domain between f_1 and f_2. These results give an indication of a role of the dissipative effects in usual cases. It is seen that those are most conspicuous near the resonance frequencies and change moderately the magnitude of G_1 in the domain between f_1 and f_2.

Figures 3–5 present the same two bubbles in glycerin. The data for glycerin were taken to be $\rho = 1260$ kg/m^3, $\eta = 1.48$ Pa s, $c = 1923$ m/s, $\sigma = 0.0657$ N/m. Figure 3 illustrates the force on the bigger bubble, bubble 1, and allows one to compare contributions to G_1 from each term of Eq. (29) for various frequency regions. It is seen from Fig. 3(a) that for low frequencies, up to about 15 kHz, the main contribution to G_1 comes from the ''acoustic streaming'' term $\mathbf{\Delta}_1$. Notice that for this frequency range the viscous wavelength, defined as $\lambda_v = 2\pi\delta_v$, is large compared with the bubble radii. For frequencies near and above the resonance frequencies of the bubbles, Fig. 3(b) and (c), the other dissipative term, $\delta_1\delta_2$, becomes dominant. Thus Fig. 3 shows that at all frequencies investigated the contribution of the ''ideal'' term Ω is negligible.

Figure 4 illustrates the force on the smaller bubble, bubble 2. The quantities G_2 and Δ_2, shown in Fig. 4(a) with negative sign, are found from Eqs. (29) to (31) by changing the subscript 1 to 2 and *vice versa*. Positive values of G_2 correspond to attraction of bubble 2 to bubble 1, and negative values to repulsion. It is seen from Fig. 4(a) that for low frequencies, where $\lambda_v \gg R_1, R_2$, the term Δ_2 is dominant and bubble 2 is repelled from bubble 1. For frequencies above 41 kHz, the term $\delta_1\delta_2$ starts to predominate, and the force on





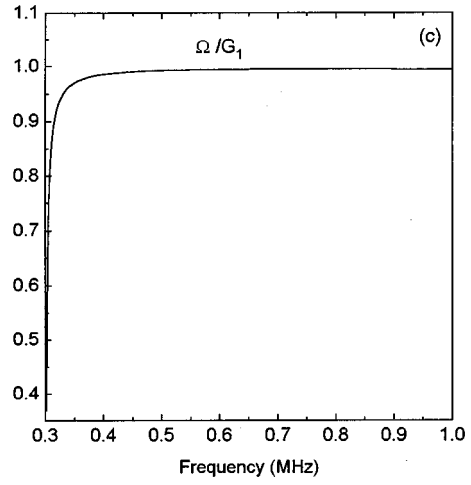
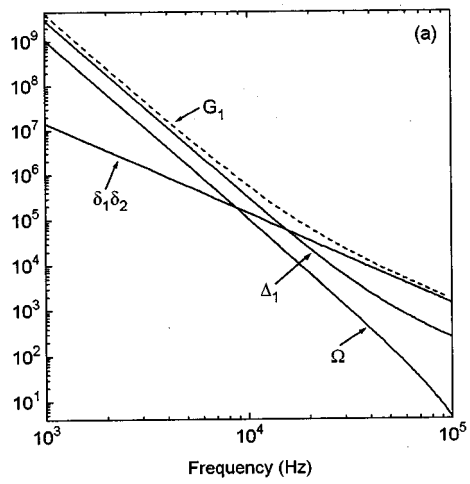
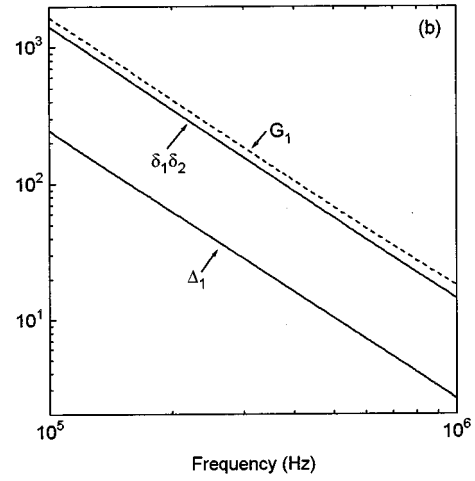


FIG. 2. Dissipative effects on Bjerknes forces between two air bubbles in water ($R_1=20\ \mu$m, $R_2=10\ \mu$m).





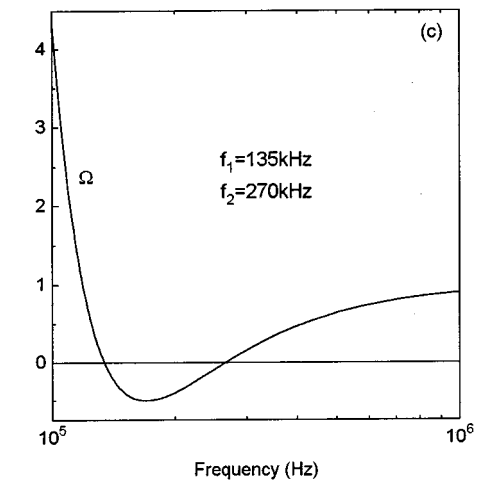
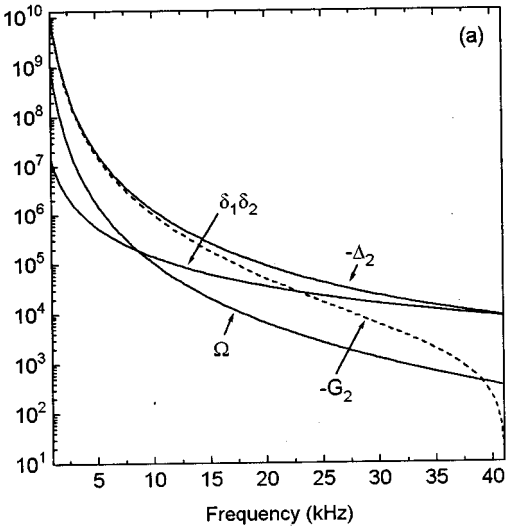


FIG. 3. Contributions from different terms of Eq. (29) to the force on the bigger bubble (bubble 1). The host liquid is glycerin, the gas inside the bubbles is air, the mean radii of the bubbles are $R_1=20\ \mu$m and $R_2=10\ \mu$m.

bubble 2 becomes attractive; see Fig. 4(b). It follows from Fig. 4(a) and comparison of Figs. 3(c) and 4(b) that the ''ideal'' term Ω is negligible in the whole frequency range, except for frequencies close to 41 kHz where $\mathbf{\Delta}_2$ and $\delta_1\delta_2$ cancel each other.

Figure 5 shows the ratio of G_2 to G_1 as a function of frequency. Note that as $G_1>0$ for all frequencies, $G_2/G_1>0$ corresponds to mutual attraction between the bubbles, while $G_2/G_1<0$ denotes that bubble 2 is repelled from



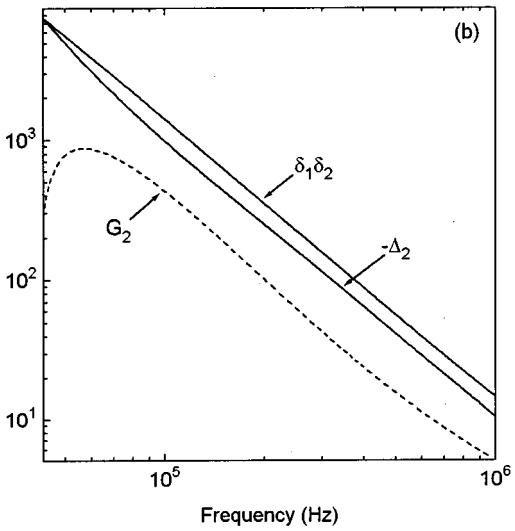


FIG. 4. Contributions from different terms to the force on the smaller bubble (bubble 2) for the same case as in Fig. 3.

bubble 1. It is seen from Fig. 5 that in general the forces on the bubbles are not equal and opposite to each other. For high frequencies, both bubbles experience attraction to each other but the force on the smaller bubble, bubble 2, is about four times smaller in magnitude than that on the bigger bubble, bubble 1. For low frequencies, the force on the smaller bubble becomes repulsive, and its magnitude is larger than that of the force on the bigger bubble.

III. CONCLUSIONS

It has been shown that dissipative effects on Bjerknes forces between two bubbles are due to two processes, namely, losses in the linear pulsations of the bubbles, which

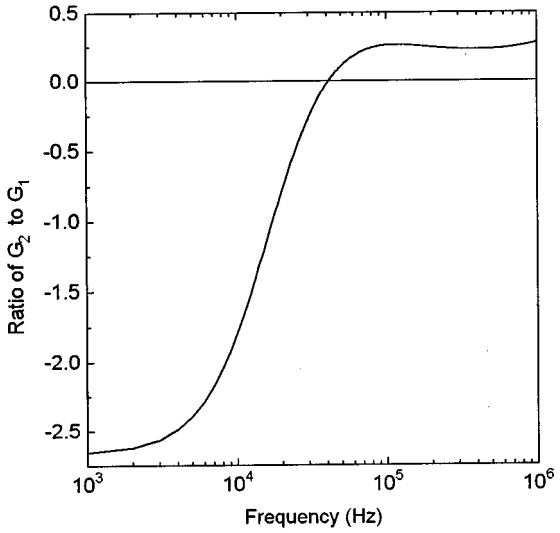


FIG. 5. Ratio of G_2 to G_1 versus frequency for the case shown in Figs. 3 and 4.

were first taken into account by Zabolotskaya,[6] and the acoustic streaming developing around the bubbles due to the viscosity of the host liquid, which has not been considered previously. It has been found that the acoustic streaming adds one more term to the Bjerknes force. This new term has been demonstrated by numerical examples to be dominant when the viscous wavelength is large compared with the bubble radii. In addition, its presence results in the forces on the bubbles which are not equal and opposite to each other.

ACKNOWLEDGMENT

This work was supported by the Ministry of Education of the Republic of Belarus.

[1] V. F. K. Bjerknes, *Fields of Force* (Columbia U. P., New York, 1906).
[2] L. A. Crum, ''Bjerknes forces on bubbles in a stationary sound field,'' J. Acoust. Soc. Am. **57**, 1363–1370 (1975).
[3] N. 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).
[4] N. 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).
[5] A. A. Doinikov and S. T. Zavtrak, ''On the mutual interaction of two gas bubbles in a sound field,'' Phys. Fluids **7**, 1923–1930 (1995).
[6] E. A. Zabolotskaya, ''Interaction of gas bubbles in a sound wave field,'' Akust. Zh. **30**, 618–623 (1984) (in Russian).
[7] L. D. Landau and E. M. Lifshitz, *Fluid Mechanics* (Pergamon, New York, 1959).
[8] A. A. Doinikov, ''Acoustic radiation pressure on a compressible sphere in a viscous fluid,'' J. Fluid Mech. **267**, 1–21 (1994).
[9] A. I. Eller, ''Damping constants of pulsating bubbles,'' J. Acoust. Soc. Am. **47**, 1469–1470 (1970).
[10] A. A. Doinikov and S. T. Zavtrak, ''Radiation effects of the interaction of two gas bubbles in a compressible nonideal fluid,'' Akust. Zh. **36**, 429–432 (1990) (in Russian).

Sound scattering from a moving turbulent medium

Zhen Ye
Institute of Ocean Sciences, Sidney, British Columbia, Canada, and Department of Physics, National Central University, Chung-li, Taiwan, Republic of China

(Received 26 June 1996; revised 12 February 1997; accepted 28 March 1997)

This paper studies sound scattering by a moving turbulent atmospheric medium, following the work of Clifford and Brown [J. Acoust. Soc. Am. **55**, 929–933 (1974)]. A formula for the sound scattering cross section is derived using the Galilean transformation and Snell's law of wave refraction. Effects of the mean flow on the scattering process are examined. Numerical examples show that, unlike the zero-flow case, there exists a nonzero cross section at right scattering angles for scattering from both velocity and temperature fluctuations. More importantly, the mean flow may give rise to significant contributions to the backscattering from turbulent eddies, in contrast to the zero-flow case and the previous results.
[S0001-4966(97)05107-2]

PACS numbers: 43.28.Fp, 43.20.Fn [LCS]

LIST OF SYMBOLS

ω	transmitted acoustic frequency
c	sound speed in the still air
ω'	acoustic frequency in the actual scattering process
$\mathbf{v}$	medium moving velocity
$\hat{v}$	$\mathbf{v}/v$—unit vector for the moving velocity
M	v/c—Mach number
$\mathbf{k}_i$	incident wave vector
$\mathbf{k}_i'$	incident wave vector in the moving medium
$\mathbf{k}_s$	scattering wave vector
$\mathbf{k}_s'$	scattering wave vector in the moving medium
$\hat{k}_i$	incident unit vector
$\hat{k}_i'$	incident unit vector in the moving medium
$\hat{k}_s$	scattering unit vector
$\hat{k}_s'$	scattering unit vector in the moving medium
k	$\lvert\mathbf{k}_i\rvert=\omega/c$
k'	$\lvert\mathbf{k}_i'\rvert=\omega'/c$
θ,ϕ	polar angles for vectors

INTRODUCTION

Over the years, there has been considerable interest in the study of acoustic scattering in turbulent media. This is not only because acoustic techniques are a powerful mean to probe various turbulent environments, such as the atmosphere, the ocean, and rivers, but also the interaction between acoustic waves and turbulent media itself is an interesting, long-standing, and theoretically challenging problem. Since the pioneering studies by Blokhintzev,[1] Kraichnan,[2] and Lighthill[3] who considered the effects of velocity fluctuation on sound scattering by turbulent media, further effort has been directed to this problem. A solution for sound scattering by turbulent media including the fluctuation of refractive index was first obtained by Batchelor[4] based on Kolmogorov's homogeneous turbulence model. This result was later rederived and improved independently by several other authors,[5] and the results are summarized in Tatarskii[6] and Ishimaru,[7] and have been reviewed by Brown and Hall.[8] The research on sound scattering from turbulence has continued to be active from both the theoretical and experimental aspects. The more recent works by, for example, Clifford and Lataitis,[9] L'Espérance *et al.*,[10] McBride *et al.*,[11] Goodman,[12] and Noble and Ostachev,[13] have shed further light on the subject.

However, most previous studies are for cases of zero mean flow. Although this may be a valid approximation for many situations, studies of effects due to nonzero mean flow are needed. The reason is twofold. First, such studies provide a justification to the approximation. Second, there are cases where the effects of the mean flow can be a major concern. In fact, how to extend the studies to scattering from moving turbulent flows has been an intriguing and important problem, not only from the viewpoint of atmospheric acoustics where the wind speed is often comparable with the sound speed in the air, but from the standpoint of ocean acoustics where the influence of flows on sound propagation and scattering poses a long-standing problem, e.g., the Doppler effect.

The early theoretical investigation of sound scattering by moving turbulence was given by Clifford and Brown[14] following the work of Brown.[15] The situation they considered can be described as follows. A horizontally stratified turbulent layer is advected by wind with a constant velocity, and is insonified by a sounder located outside the layer. The scattered signal is then recorded by a receiver which is also located outside the layer. The medium outside the layer is assumed to be stationary. Using Snell's law for sound refraction, waves inside and outside the turbulent medium can be related at the interface. Clifford and Brown derived a formula for the sound scattering cross section of the moving turbulent medium. In their approach, it was shown that except for a modified angular dependence, basically there is no change in the form of scattering cross section. Their results

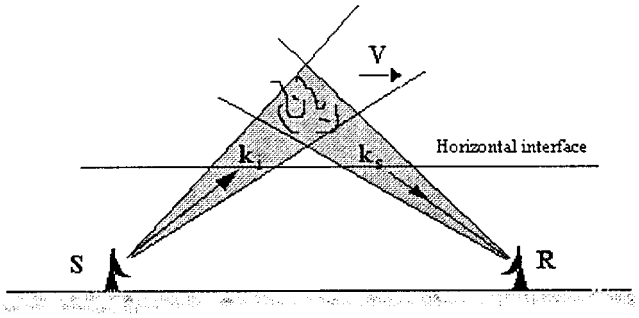


FIG. 1. A conceptual drawing of the acoustic scattering process. Shown here is the sounder S and receiver R. The turbulent eddies are advected by the medium with a velocity **v**.

show that there are several important influences of the nonzero mean flow on the turbulent scattering process, particularly at scattering angles where the zero mean flow turbulence cross section is small. The scattering pattern can thus be modified significantly.

In this article, following the spirit of Clifford and Brown, we reconsider sound scattering from a moving turbulent medium. The situation considered here is identical to Clifford and Brown's. Using the Galilean transformation, we first write the incident wave in the frame that is comoving with the medium. Second, the refracted and incident waves are connected in the comoving frame by Snell's law of refraction that requires the phase to be continuous across the boundary. We will then derive a formula for the sound scattering cross section for a moving turbulent medium as a function of scattering angles. It will be shown that this scattering cross section can be written in terms of the scattering function that would be obtained without mean flow. The effects of mean medium movement on the scattering cross section will be examined by several examples in the context of Kolmogorov's isotropic turbulence theory. We will point out that due to the advection, there can be a significant contribution to the backscattering cross section from velocity fluctuation, in contrast to the zero mean flow case, and also different from the result of Clifford and Brown.

I. ANALYSIS

The scattering problem considered in this article is similar to Ref. 14, and is illustrated in Fig. 1. Here is shown a turbulent variability embedded in a stratified medium advected by a mean flow with velocity **v** assumed to be known. Note that although the mean flow direction is shown in the figure to be parallel to the interface (not necessarily restricted to the plane spanned by $\mathbf{k}_i$ and $\mathbf{k}_s$, i.e., the reception plane), the later development is also valid when the direction is not parallel to the interface, which may be the case when the turbulent medium is advected as a whole. In Ref. 14, Clifford and Brown actually considered the more general case when the mean flow direction is at an oblique angle β from the interface.

We assume that the medium around the acoustic system stays still. The advected medium is insonified by a sounder S from the direction of $\hat{k}_i$. The acoustic frequency transmitted at S is ω and, for simplicity, the magnitude of **v** is taken to be constant. The echo returns along the direction of $\hat{k}_s$ and reaches the receiver (R). Here, $\hat{k}_i$ and $\hat{k}_s$ denote the transmitting and receiving unit vectors, respectively, i.e., $|\hat{k}_i| = |\hat{k}_s| = 1$. We note that the three vectors **v**, $\hat{k}_i$, and $\hat{k}_s$ are not necessarily in the same plane. The scattering volume is defined as the intersection of the incident and scattered beams.

The interface between the advected medium and the stationary medium is assumed to be horizontal, as we can always mathematically adjust the incidence direction to take into account the case of nonhorizontal interface. We set up two rectangular frames: one is associated with the acoustic systems $\mathscr{K}$ (stationary ground frame), and the other is moving with the turbulent medium $\mathscr{K}'$ (comoving). For convenience we assume the axes of the two coordinate systems to be parallel to each other, respectively, with the z axis being upward. Clearly, two steps can be invoked to solve for the sound scattering cross section as observed by the acoustic receiving system. First, the scattering cross section is calculated in the frame moving with the moving medium. Then the scattering cross section to be observed is obtained by a proper transformation. For convenience, in the following all the quantities in frame $\mathscr{K}'$ are primed. In addition, all quantities with a hat denote the unit vectors.

A. Scattering cross section in the comoving system

First we consider the scattering in the frame moving with the medium ($\mathscr{K}'$). Morse and Ingard (Chap. 11 in Ref. 16) shows that the wave equation in this frame is identical to that when the medium is at rest.

The Born approximation leads to the wave equation near the scattering region in $\mathscr{K}'$[12,17]

$$(\nabla'^2 + k'^2)p = \frac{1}{\rho}\,\boldsymbol{\nabla}'\rho\cdot\boldsymbol{\nabla}' p_0 - 2\,\frac{\delta c}{c}\,\nabla'^2 p_0 - 2\boldsymbol{\nabla}'\cdot\left(\hat{k}_i\cdot\frac{\delta\mathbf{u}}{c}\,\boldsymbol{\nabla}' p_0\right), \tag{1}$$

where ρ is the mass density of the turbulent medium, δc is the propagation speed fluctuation caused by, for example, the temperature variability, and $\delta\mathbf{u}$ is the turbulence velocity fluctuation. Moreover, in the above

$$k' = \frac{\omega'}{c}, \tag{2}$$

with ω' being the acoustic frequency in the comoving system to be determined later. Here, p_0 denotes the incident wave impinging on the turbulence, and is given by

$$p_0 = A(r')e^{ik'\hat{k}'_r\cdot\mathbf{r}' - i\omega' t'}. \tag{3}$$

The change of coordinates from the comoving system $\mathscr{K}'$ to the observation system $\mathscr{K}$ is given by the Galilean transformation[18]

$$\mathbf{r} = \mathbf{r} + \mathbf{v}t, \quad t = t'. \tag{4}$$

The corresponding spatial and temporal derivatives are given by

$$\boldsymbol{\nabla} = \boldsymbol{\nabla}', \quad \frac{\partial}{\partial t} = \frac{\partial}{\partial t'} - \mathbf{v}\cdot\boldsymbol{\nabla}'. \tag{5}$$

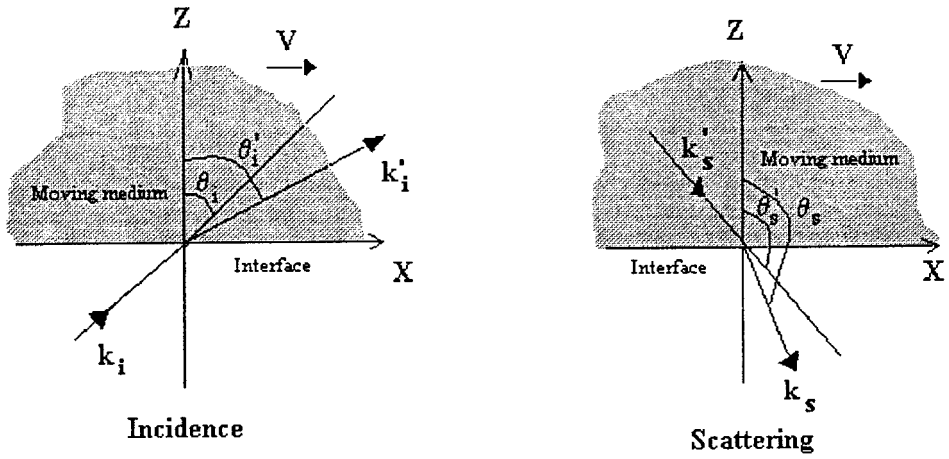


FIG. 2. Wave refraction at the boundary. Here we consider scattering in the x-z plane.

In deriving Eq. (1) and in what follows, the usual far-field condition is used. We further assume the turbulence to be statistically homogeneous and isotropic.

From the wave equation in (1) the scattering cross section per unit volume per unit solid angle in the moving coordinates can be readily calculated using the procedure outlined in Tatarskii,[6] and Goodman and Kemp.[12] They express the scattering cross section in the atmosphere as

$$\sigma'(\hat{k}_i', \hat{k}_s') = \frac{\pi}{2} k'^4 (\hat{k}_i' \cdot \hat{k}_s')^2 \left[\frac{\Phi_T(K')}{T_0^2} + \frac{(1+\hat{k}_i' \cdot \hat{k}_s') E(K')}{2\pi c^2 K'^2} \right], \tag{6}$$

where $K'^2 = 2k'^2(1-\hat{k}_i' \cdot \hat{k}_s')$, $\hat{k}_s'$ is the unit vector denoting the scattering direction, Φ_T is the correlation spectrum for the temperature fluctuation, E is the energy spectral density of the turbulence, and T_0 is the average temperature.

B. Scattering cross section as observed from the ground system

The wave transmitted from the ground must pass through the interface between the moving and stationary media, and then is scattered by the turbulent variability. The scattered wave will pass across the boundary to reach the receiver. The process is shown in Fig. 2. The waves inside and outside the moving medium can be connected by Snell's law of wave refraction. Snell's law of refraction takes a special form at the interface between a fluid at rest and a moving fluid. We apply Snell's law in the comoving system.

First transform all quantities into the moving frame. Consider a unit incident wave described by

$$p_{\text{in}} = \exp(ik\hat{k}_i \cdot \mathbf{r} - i\omega t), \tag{7}$$

where ω is the transmitted acoustic frequency, $k = \omega/c$. Using the Galilean transformation in Eq. (4), the incident wave can be written in the comoving system as

$$p_{\text{in}}' = \exp(ik\hat{k}_i \cdot \mathbf{r}' - i\omega' t'). \tag{8}$$

In this frame, the acoustic frequency is found as

$$\omega' = \omega(1 - M\hat{k}_i \cdot \hat{v}), \tag{9}$$

with $M = v/c$, the Mach number, and $\hat{v} = \mathbf{v}/v$, the unit factor in the direction of the velocity.

Since the acoustic frequency is invariant when the wave propagates across the boundary, the refracted wave can be written as

$$p_r' \sim \exp(ik'\hat{k}_i' \cdot \mathbf{r}' - i\omega' t), \tag{10}$$

where $\hat{k}_i'$ is the phase normal of the refracted wave, which is invariant under the Galilean transformation, and the wave number in the moving medium, k', is found to be

$$k' = \frac{\omega'}{c}, \tag{11}$$

which enters into Eq. (6). At the interface, Snell's law states that

$$k\hat{k}_i \cdot \hat{n} = k'\hat{k}_i' \cdot \hat{n}, \tag{12}$$

where $\hat{n}$ is an arbitrary unit vector parallel to the boundary surface.

Using the polar coordinates, the incident and refracted unit vectors can be conveniently written as

$$\begin{aligned} \hat{k}_i &= (\sin\theta_i \cos\phi_i,\ \sin\theta_i \sin\phi_i,\ \cos\theta_i), \\ \hat{k}_i' &= (\sin\theta_i' \cos\phi_i',\ \sin\theta_i' \sin\phi_i',\ \cos\theta_i'), \end{aligned} \tag{13}$$

where θ is the angle relative to the upward direction, and ϕ is the azimuthal angle. The same expressions also hold for the scattering direction to be encountered later.

In the polar coordinates, the boundary condition in Eq. (12) leads to

$$\phi_i' = \phi_i, \tag{14}$$

and

$$\sin\theta_i' = \frac{\sin\theta_i}{(1 - M\hat{k}_i \cdot \hat{v})}. \tag{15}$$

This is generally different from the boundary condition implied by Eq. (14) in Ref. 14. For comparison, the result in this reference is written out as

$$\sin\theta_i' = \sin\theta_i(1 + M\hat{k}_i' \cdot \hat{v}). \tag{16}$$

However, Eq. (16) will reduce to Eq. (15) when **v** is parallel to the interface. Note that in Ref. 14, **v** is not parallel to the interface; it makes an angle β from the interface, as suggested by their Eq. (12) and Fig. 2. The critical angle for total reflection indicated by Eq. (15) is determined by $\sin\theta_i = 1 - M\hat{k}_i \cdot \hat{v}$, which can be solved numerically. Some special cases for Eq. (15) can be found in Refs. 16, 19, and 20.

The similar relations can be found for the scattered wave. Again, it is convenient to write the scattering unit vectors in terms of the polar coordinates

$$\begin{aligned} \hat{k}_s &= (\sin\theta_s \cos\phi_s,\ \sin\theta_s \sin\phi_s,\ \cos\theta_s), \\ \hat{k}_s' &= (\sin\theta_s' \cos\phi_s',\ \sin\theta_s' \sin\phi_s',\ \cos\theta_s'). \end{aligned} \tag{17}$$

Since the acoustic wavelength does not change in the scattering process, we have

$$|\mathbf{k}_s'| = |\mathbf{k}_i'| = k'.$$

Under the Galilean transformation, we further obtain

$$|\mathbf{k}_s| = k_s = \frac{k'}{1 - M\hat{k}_s \cdot \hat{v}}.$$

The following boundary condition for the scattered wave

$$k'\hat{k}_s' \cdot \hat{n} = k_s \hat{k}_s \cdot \hat{n}, \tag{18}$$

leads to

$$\phi_s' = \phi_s, \tag{19}$$

and

$$\sin\theta_s' = \frac{\sin\theta_s}{1 - M\hat{k}_s \cdot \hat{v}}. \tag{20}$$

Clearly, θ_i', ϕ_i' and θ_s', ϕ_s' are true angles for the incidence and scattering, whereas θ_i, ϕ_i and θ_s, ϕ_s are apparent angles observed in the ground system. They are related through Eqs. (14), (15), (19), and (20).

To find the scattering cross section from Eq. (6), we need to find the quantity $\hat{k}_i' \cdot \hat{k}_s'$ in terms of $\hat{k}_i$ and $\hat{k}_s$. Using Eqs. (14), (15), (19), and (20), we have

$$\begin{aligned} \hat{k}_i' \cdot \hat{k}_s' = {} & \frac{1}{(1 - M\hat{k}_i \cdot \hat{v})(1 - M\hat{k}_s \cdot \hat{v})} \\ & \times [\sin\theta_i \sin\theta_s \cos(\phi_s - \phi_i) \\ & - \sqrt{[(1 - M\hat{k}_i \cdot \hat{v})^2 - \sin^2\theta_i]} \\ & \times \sqrt{[(1 - M\hat{k}_s \cdot \hat{v})^2 - \sin^2\theta_s]}, \end{aligned} \tag{21}$$

where

$$\hat{k}_{i,s} \cdot \hat{v} = \sin\theta_{i,s} \sin\theta_v \cos(\phi_{i,s} - \phi_v) + \cos\theta_{i,s} \cos\theta_v, \tag{22}$$

in which θ_v, ϕ_v are polar angles for the unit vector $\hat{v}$. According to Ref. 14, the scattering cross section as observed in the ground frame is obtained by substituting the above expression into Eq. (6). This implicitly assumes that the waves refracted by low speed wind with no significant change in amplitude. For example, as pointed out by a referee, with $M = 0.05$ and $\theta_i = 30$ deg, the amplitude is only reduced by a small fraction 0.008. From Eq. (21), it is clear that the quantity $\hat{k}_i' \cdot \hat{k}_s'$ can be solved without resorting to approximations.

We note another ambiguity in Clifford and Brown.[14] According to Eq. (25) in this reference, the backscattering in the ground observation would imply that the scattering in the moving medium is also backward. This can be seen by substituting $\theta = \pi$ into their Eq. (25), which leads to $\hat{\mathbf{e}}_0 \cdot \hat{\mathbf{e}}_1 = -1$, with $\hat{\mathbf{e}}_0$ and $\hat{\mathbf{e}}_1$ being the unit vectors for the incidence and scattering directions in the moving medium, respectively. This is in conflict to the nonreciprocity for wave refraction in moving media (referring to, e.g., Fig. 11.5 and the related text in Morse and Ingard[16]).

II. DISCUSSION

In this section we consider a few examples to examine the effects due to the mean flow. Although the above results are valid for any scattering scenario, without losing generality we focus our attention on a special case, namely, the direction of the flow is parallel to the plane spanned by the transmitting direction $\hat{k}_i$ and the receiving direction $\hat{k}_s$ (reception plane). For convenience we assume that the plane is the x-z plane. In this case, we have $\phi_i = 0$, and ϕ_s can be either 0 or ϕ depends on the sign of the x component of $\hat{k}_s$. Furthermore, we take the conventional assumption that the turbulence can be described by the Kolmogorov isotropic spectrum for $\Phi_T(K)$ and $E(K)$:[7]

$$\Phi_T(K) = 0.033 C_T^2 K^{-11/3}, \tag{23}$$

and

$$E(K) = 0.76 C^2 \epsilon^{2/3} K^{-5/3}, \tag{24}$$

where C_T is the structure constant of the temperature fluctuation, ϵ is the rate of dissipation of turbulent energy, $C^2 \approx 1.9$ is a numerical constant. Equations (23) and (24) are valid for wave number K within the inertial subrange.

With Eqs. (23) and (24), the scattering cross sections from Eq. (6) can be written in the following form:

$$\begin{aligned} \sigma(\hat{k}_i, \hat{k}_s) = {} & A(\omega, T_0, C_T^2)(1 - M\hat{k}_i \cdot \hat{v})^{1/3}(\hat{k}_i' \cdot \hat{k}_s')^2 \\ & \times (1 - \hat{k}_i' \cdot \hat{k}_s')^{-11/6} + B(\omega, \epsilon, C)(1 - M\hat{k}_i \cdot \hat{v})^{1/3} \\ & \times (\hat{k}_i' \cdot \hat{k}_s')^2 (1 + \hat{k}_i' \cdot \hat{k}_s')(1 - \hat{k}_i' \cdot \hat{k}_s')^{-11/6}, \end{aligned} \tag{25}$$

for the atmosphere turbulence, where $\hat{k}_i' \cdot \hat{k}_s'$ is related to the apparent incidence and scattering vectors through Eq. (21). In the above equation, we have

$$A(\omega, T_0, C_T^2) = 0.0145 \frac{C_T^2}{T_0^2} \left(\frac{\omega}{c}\right)^{1/3},$$

$$B(\omega, \epsilon, C) = 0.053 \frac{C^2 \epsilon^{2/3}}{c^2} \left(\frac{\omega}{c}\right)^{1/3}.$$

Now we consider the influences of the mean flow on the scattering process. We choose $M = 0.05$ or $v = 17$ m/s as an example. In order to take into account the case of nonhorizontal interface, we consider two incident angles, i.e., θ_i

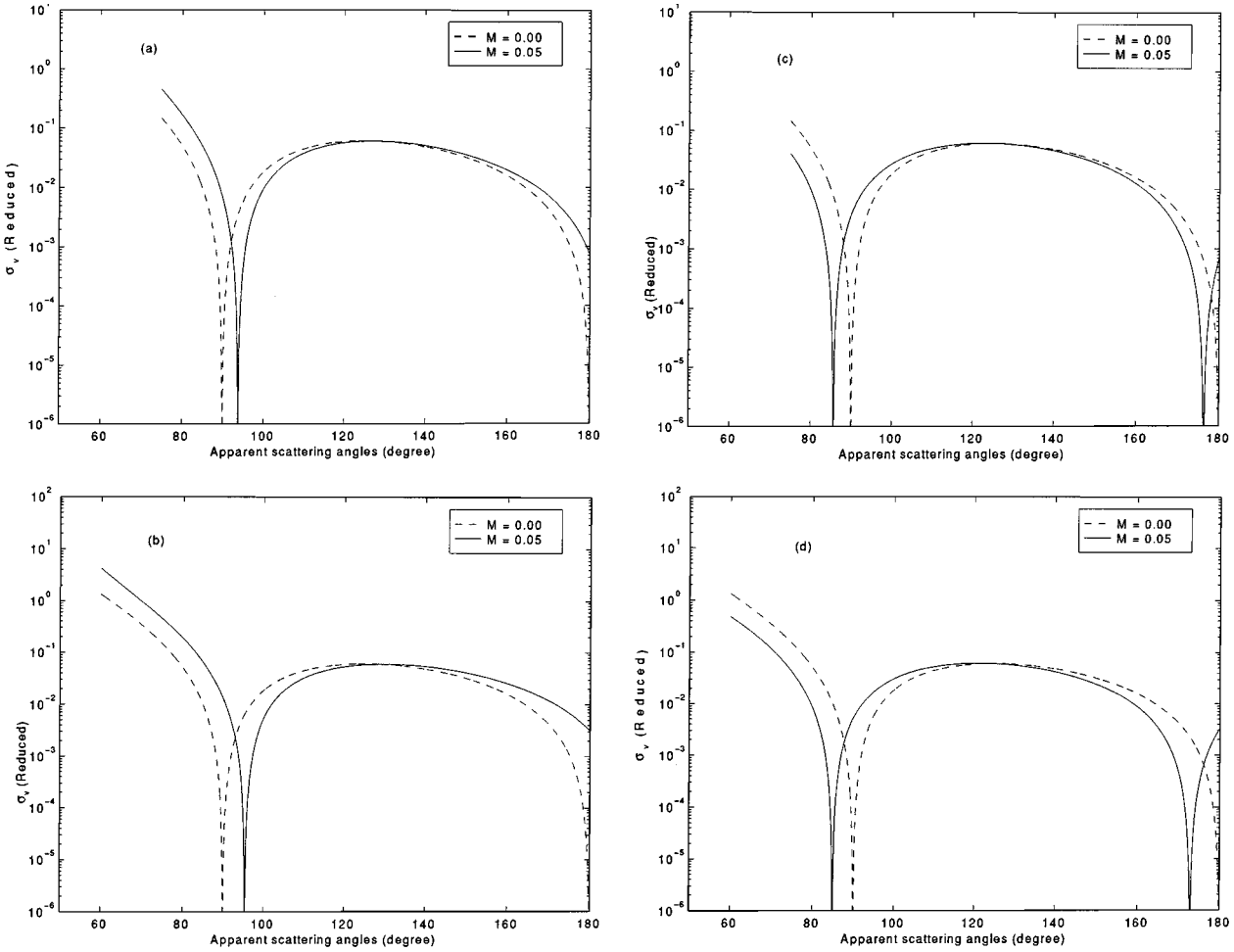


FIG. 3. Normalized scattering cross section σ_v due to turbulence fluctuation as a function of apparent scattering angle at $M=0$, 0.05: (a) $\theta_i=\pi/4$, $\phi_i=0$, and $\theta_v=\pi/2$, $\phi_v=0$, i.e., the horizontal wind in the positive x-axis direction; (b) $\theta_i=\pi/3$, $\phi_i=0$, and $\theta_v=\pi/2$, $\phi_v=0$; (c) $\theta_i=\pi/4$, $\phi_i=0$, and $\theta_v=\pi/2$, $\phi_v=\pi$, i.e., the horizontal wind in the negative x-axis direction; (d) $\theta_i=\pi/3$, $\phi_i=0$, and $\theta_v=\pi/2$, $\phi_v=\pi$. The scattering cross section is normalized by the constant $B(\omega,\epsilon,C)$.

$=\pi/4$ and $\pi/3$. Here the sound speed in the air is assumed to be 340 m/s. In Fig. 3 we plot the normalized scattering cross section due to velocity fluctuation (σ_v) as a function of apparent scattering angle ($\theta=\theta_s-\theta_i$) for four cases. Case (a) corresponds to the case considered in Ref. 14. In these cases, the scattering direction is represented by (θ_s,ϕ_s), and ϕ_s equals 0 (π) when the x component of $\hat{k}_s$ is positive (negative). Figure 3 shows a few interesting aspects of the effects of the mean flow on the scattering process.

(1) Without mean flow, the scattering due to velocity fluctuation produces nulls at the right angles ($\theta=\pi/2$) and the backscattering angle ($\theta=\pi$) to the incident direction.

(2) The figure shows that the most important influences of the mean flow are represented by the modifications of the scattering pattern.

(3) Although these influences are small for most of the scattering angles, they take on great significance at angles where the zero mean flow scattering cross section is small; it is similar to what has been found by Clifford and Brown.[14]

(4) In general, the null positions in the scattering cross section are shifted due to the nonzero mean flow; consequently, there will be nonvanishing scattering cross section at right scattering angles, unlike the zero mean flow case.

(5) The null positions can be shifted to either a larger or a smaller angle, depending on the wind direction; when the wind blows against the sounding direction, there are two null positions in the scattering cross section: one at an angle smaller than $\pi/2$, and the other at an angle smaller than π [Fig. 3(c) and (d)].

(6) A striking feature associated with the scattering from the velocity fluctuation is that the backscattering ($\theta=\pi$) no longer vanishes, in contrast to the zero mean flow case, and also different from the consideration of Clifford and Brown,[14] who showed that the nonzero mean flow has no effect on the backscattering (refer to Fig. 4 in Ref. 14). The occurrence of this feature is because the apparent backscattering does not appear to be strictly backscattering in the moving medium; that is equivalent to say, if a wave that has been transmitted from the still medium to the moving medium is reversed in direction, it will not follow along the path of the initial ray.

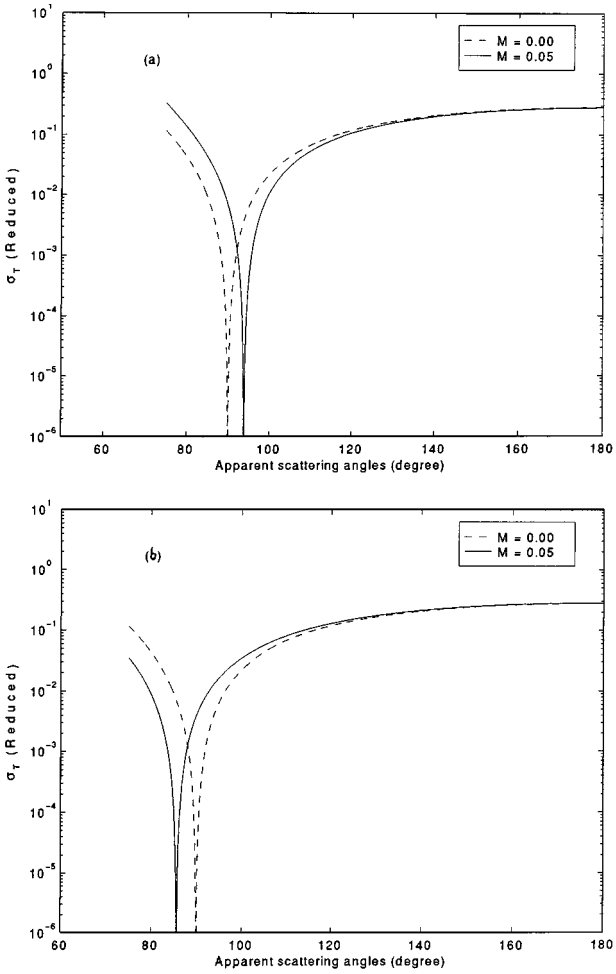


FIG. 4. Normalized scattering cross section σ_T due to temperature fluctuation as a function of apparent scattering angle at $M=0$, 0.05: (a) $\theta_i=\pi/4$, $\phi_i=0$, and $\theta_v=\pi/2$, $\phi_v=0$, i.e., the horizontal wind in the positive x-axis direction; (b) $\theta_i=\pi/4$, $\phi_i=0$, and $\theta_v=\pi/2$, $\phi_v=\pi$, i.e., the horizontal wind in the negative x-axis direction; The scattering cross section is normalized by the constant $A(\omega,T_0,C_T^2)$.

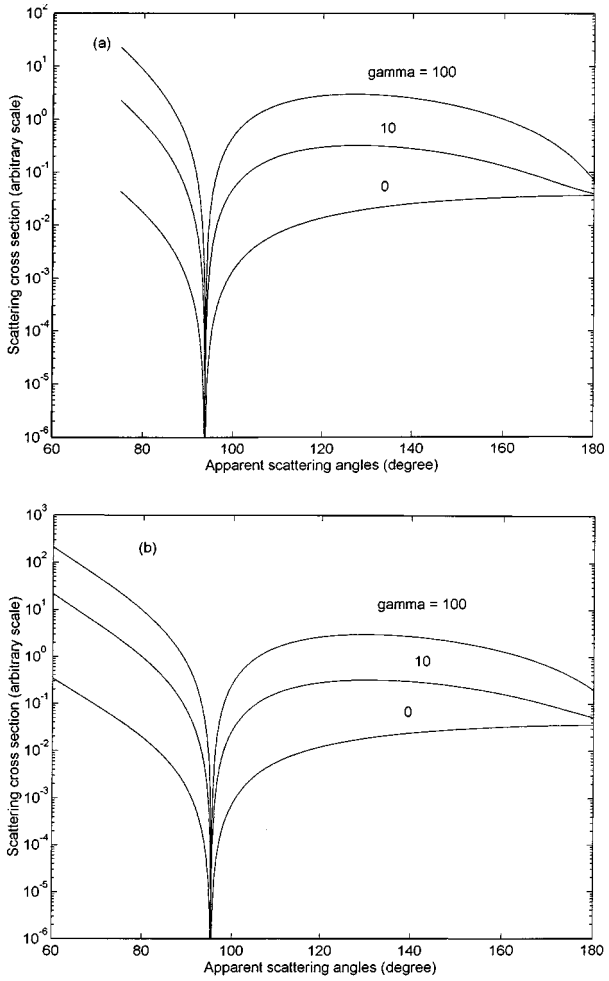


FIG. 5. Total scattering function as a function of apparent scattering angle at $M=0.05$ for various γ values: (a) the incident angle is taken as $\theta_i=\pi/4$, $\phi_i=0$, while in (b) we take $\theta_i=\pi/3$, $\phi_i=0$.

(7) The backscattering due to the velocity fluctuation can be rather significant, especially for larger θ_i, i.e., larger slanting angles.
(8) Other than near backscattering, the feature of the effects of the mean flow is similar to what has been described in Clifford and Brown.[8]

Figure 4 shows the normalized scattering cross section due to temperature fluctuation (σ_T) as a function of apparent scattering angle for two cases considered above. Here we see again that (1) the nonzero flow motion modifies the scattering pattern; (2) the scattering cross section no longer vanishes at right angles to the incident direction; (3) at near backscattering angles, the effect of nonzero mean flow is little; (4) the results in Fig. 4 are very similar to that obtained in Ref. 14; (5) the null positions in the scattering cross section due to the temperature fluctuation are located at one of the null angles of that due to the velocity fluctuation.

We also plot the total scattering cross section as a function of apparent scattering angle. As usual, we define a parameter γ as[8]

$$\gamma=\frac{C^2\epsilon^{2/3}T_0^2}{C_T^2c^2} \tag{26}$$

to denote the strength ratio of the velocity and temperature fluctuations. In Fig. 5 we plot the total scattering cross section against the apparent scattering angle[8] for three typical γ values: $\gamma=0$, 10, and 100. The wind is assumed to be horizontal along the x axis, and $M=0.05$. In (a), the incident angle is taken as $\theta_i=\pi/4$, $\phi_i=0$, while in (b) we take $\theta_i=\pi/3$, $\phi_i=0$. Here, we can see clearly that the backscattering can be significantly enhanced by the scattering from the velocity fluctuation. In particular, in case (a) the backscattering is enhanced by 10% for $\gamma=10$ and the enhancement becomes 100% for $\gamma=100$. In case (b) the backscattering due to the velocity fluctuation takes about 33% of the total backscattering, and this portion increases by ten times when γ goes up to 100.

III. SUMMARY

In this article we considered sound scattering from a moving turbulent medium. An analytic formula for the sound scattering cross section was derived using the Galilean transformation and Snell's law of wave refraction. The effects of

the nonzero mean flow on scattering by the velocity and temperature fluctuations were examined in detail. Although only the atmospheric case was considered, the theoretical derivation may be extended to sound scattering from other moving media as well. It is shown that the nonzero mean flow can have significant effects on the scattering at certain scattering angles, particularly on the scattering due to the velocity fluctuation. A major effect of the nonzero mean flow is that there may exist significant nonvanishing scattering due to the velocity fluctuation at both the backscattering angle and at right scattering angles, in contrast to the zero mean flow case. This is a rather important point to be recognized when performing experiments to resolve the scattering by temperature fluctuation from the scattering by velocity fluctuation, since it has been usually assumed that there is no backscattering from velocity fluctuation so that the temperature variation may be inferred by backscattering. Finally, we note that there is in reality no sharp interface, as the buildup of wind above a ground is gradual. This complicated case is not considered here.

ACKNOWLEDGMENTS

The associate editor, L. C. Sutherland, and the three referees are thanked for useful constructive suggestions and useful comments, which are very helpful for preparing the final version of the manuscript.

[1] D. Blokhintzev, ''The propagation of sound in an inhomogeneous and moving medium I, II,'' J. Acoust. Soc. Am. **18**, 322–334 (1946).
[2] R. H. Kraichnan, ''The scattering of sound in a turbulent medium,'' J. Acoust. Soc. Am. **25**, 1096–1104 (1953).
[3] M. J. Lighthill, ''On the energy scattered from the interaction of turbulence with sound or shock waves,'' Proc. Camb. Philos. Soc. **49**, 531–555 (1953).
[4] G. K. Batchelor, ''Wave scattering due to turbulence,'' in *Symposium on Naval Hydraulics*, edited by F. S. Sherman (National Academy of Sciences, Washington, DC, 1957), pp. 409–423.
[5] For example, M. A. Kallistratova and V. I. Tatarskii, ''Accounting for wind turbulence in the calculation of sound scattering in the atmosphere,'' Sov. Phys. Acoust. **6**, 503–505 (1960); A. S. Monin, ''Characteristics of the scattering of sound in a turbulent atmosphere,'' Sov. Phys. Acoust. **7**, 370–373 (1962); and Ref. 12.
[6] V. I. Tatarskii, ''The effects of the turbulent atmosphere on wave propagation,'' U.S. Dept. of Commerce, TT-68-50464, Springfield, Virginia (1971).
[7] A. Ishimaru, *Wave Propagation and Scattering in Random Medium* (Academic, New York, 1978), Vols. I and II.
[8] E. H. Brown and F. F. Hall, ''Advances in atmospheric acoustics,'' Rev. Geophys. Space Phys. **16**, 47–110 (1978).
[9] S. F. Clifford and R. Lataitis, ''Turbulence effects on acoustic wave propagation over a smooth surface,'' J. Acoust. Soc. Am. **73**, 1545–1550 (1983).
[10] A. L'Espérance, Y. Gabillet, and G. A. Daigle, ''Outdoor sound propagation in the presence of atmospheric turbulence,'' J. Acoust. Soc. Am. **98**, 570–579 (1995).
[11] W. E. McBride, H. E. Bass, R. Raspet, and K. Gilbert, ''Scattering of sound by atmospheric turbulence,'' J. Acoust. Soc. Am. **90**, 3314–3325 (1991).
[12] L. Goodman and K. A. Kemp, ''Scattering from volume variability,'' J. Geophys. Res. **86** (C5), 4083–4088 (1981).
[13] J. M. Noble and V. E. Ostachev, ''New formulations for the scattering of sound in a moving random medium,'' J. Acoust. Soc. Am. **98**, 2924 (1995) (A).
[14] S. F. Clifford and E. H. Brown, ''Acoustic scattering from a moving turbulent medium,'' J. Acoust. Soc. Am. **55**, 929–933 (1974).
[15] E. H. Brown, ''Acoustic-Doppler radar scattering equation and general solution,'' J. Acoust. Soc. Am. **52**, 1391–1396 (1972).
[16] P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (Princeton U. P., Princeton, NJ, 1986).
[17] L. Goodman, ''Acoustic scattering from ocean microstructure,'' J. Geophys. Res. **95** (C7), 11 557–11 573 (1990).
[18] J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1975).
[19] J. Miles, ''On the reflection of sound at an interface of relative motion,'' J. Acoust. Soc. Am. **29**, 226–228 (1957).
[20] H. S. Ribner, ''Reflection, transmission, and amplification of sound by a moving medium,'' J. Acoust. Soc. Am. **29**, 435–441 (1957).

Acoustic scattering by atmospheric turbules

George H. Goedecke
Department of Physics, New Mexico State University, Las Cruces, New Mexico 88003-0001

Harry J. Auvermann
U.S. Army Research Laboratory, Adelphi, Maryland 20783-1155

(Received 15 June 1996; revised 3 February 1997; accepted 28 April 1997)

Atmospheric turbulence is modeled as a collection of self-similar localized eddies, called turbules. Turbulent temperature variation and solenoidal velocity structure function spectra and the corresponding average acoustic scattering cross sections are calculated for several isotropic homogeneous turbule ensembles. Different scaling laws for turbule strengths, number densities, and sizes produce different power-law spectra independent of turbule morphology in an ''inertial range'' of the spectral variable K. For fractal size scaling and Kolmogorov power law $\propto K^{-11/3}$ in the inertial range, not only do turbule strengths scale like the one-third power of the size, but also the turbule packing fractions are scale invariant, as are the expressions derived for the structure parameters (C_T^2, C_v^2). The inertial range boundaries of the spectral variable and scattering angles are easily estimated from the inner and outer scales of the turbulence. They depend weakly on turbule morphology, while the spectra and cross sections outside the inertial ranges depend strongly on it. Scattering at angles outside the inertial range, which occurs in practical cases, is much weaker than predicted by the Kolmogorov spectrum. For Gaussian turbule ensembles, quasianalytic forms are obtained for the spectra and scattering cross sections and for the structure functions themselves.
 [S0001-4966(97)04108-8]

PACS numbers: 43.28.Py, 43.20.Fn [LCS]

INTRODUCTION

In both classical and quantum physics, scattering is very often described as scattering of waves by one or more particles. The particles may be the ''fundamental'' ones with which the waves interact, or they may be ''quasiparticles'' described by an effective interaction with the wave field and possibly composed of billions of the fundamental scattering particles. For example, for many electromagnetic waves, electrons and nuclei or ions may be taken as the fundamental scattering particles, while atmospheric aerosol particles are quasiparticles, described by their refractive index morphology for wavelengths much greater than the average atomic spacing in an aerosol particle.

If the properties and distributions of the particles that constitute a random medium are specified, then the statistical functions that determine wave propagation and scattering in the medium can be derived in detail, even for anisotropic and/or inhomogeneous cases. However, if a particle substructure is not used, then the general behavior of the relevant statistical functions can often be inferred by other means.

Atmospheric turbulence is most often treated as a stationary random medium for scattering and propagation of both electromagnetic and acoustic waves.[1,2] The statistical functions that determine the ensemble average acoustic scattering in Born approximation include the spectra of the temperature variation and velocity structure functions of the turbulence, among others. Conventionally, the general behavior of each of these spectra is inferred using symmetry arguments, dimensional analysis, the fluid equations, and expected asymptotic behavior;[1–4] their detailed behavior and values cannot be derived. Even for isotropic homogeneous turbulence, several different spectra are in common use; for example, the Kolmogorov and von Karman spectra[1,2] and generalizations and modifications by Tatarski,[1] Monin and Yaglom,[5] Hill,[6] and Ostashev *et al.*[7]

There are no ''fundamental'' particles with which sound waves interact. However, atmospheric turbulence has long been regarded as composed of eddies of many different sizes. Implicit in the eddy concept is the notion of localization, i.e., an individual turbulent eddy involves fluctuations that go rapidly to zero beyond some distance or ''scale length'' from its nominal center. This localization makes it a suitable quasiparticle. Localized turbulent temperature eddies, designated ''turbules,'' have already been used.[8,9] In this paper, we extend the notion of turbules to include localized solenoidal turbulent flows as well, and we regard the turbulence as a superposition of randomly located self-similar turbules having sizes ranging between designated inner and outer scale lengths.

In later work, we plan to utilize a turbule model to achieve our main goal of predicting observed acoustic scattering and its fluctuations by realistic anisotropic inhomogeneous turbulence. In this paper, we focus on a nontrivial and essential first step: To ascertain whether a turbule model of isotropic homogeneous turbulence yields physically reasonable structure function spectra that can be matched to the Kolmogorov spectra in a so-called ''inertial range,'' independently of turbule morphology. Also, we expect the model to yield expressions for the boundaries of the inertial range and the behavior of the scattering both inside and outside the corresponding inertial range of scattering angles.

In Sec. I, we give a brief reprise of the conventional

treatment of acoustic scattering by isotropic homogeneous atmospheric turbulence. In Sec. II, we write down scalable temperature variation and velocity potential functions and scattering amplitudes for a single model turbule. In Sec. III, we derive general expressions for temperature and velocity spectra and scattering efficiencies for an isotropic ensemble of one-turbule systems, and we provide expressions and graphs for two examples of comparable size but different morphology. We show that an ensemble with members consisting of a spherically symmetric nonuniformly rotating turbule with randomly directed rotation axis may replace a general isotropic ensemble of one-turbule systems, and we derive the minimal azimuthal dependence of the velocity scattering for a given ensemble member. In Sec. IV, we calculate the incoherent scattering produced by an isotropic homogeneous ensemble whose members contain self-similar turbules of many sizes distributed uniformly in space. We assume that turbule number densities and their temperature and velocity amplitudes scale with size, and that sizes also scale according to a power law in the size index, from the outer scale to the inner. Then we show that this turbule model of isotropic homogeneous turbulence indeed accomplishes what was conjectured above, and allows some detailed predictions that cannot be achieved by conventional treatments. We include graphs of typical example structure functions and their spectra; of the dependence of the inertial range of scattering angles on outer and inner scale size parameters; and of the differential scattering efficiencies for some of the example spectra. In Sec. V, we provide a detailed list of results and a short discussion of future work needed in order to predict acoustic scattering by anisotropic inhomogeneous turbulence using turbule models.

I. SCATTERING BY TURBULENCE

A. Scattering amplitudes

The conventional Born approximation far-field scattering amplitude $f(\hat{\mathbf{r}})$ (dimension, length) for scattering of an incident acoustic plane wave of angular frequency ω, wavelength λ by stationary atmospheric turbulence is given by the Monin[10] results

$$f(\hat{\mathbf{r}})=f_T(\hat{\mathbf{r}})+f_v(\hat{\mathbf{r}}), \tag{1}$$

where

$$\begin{aligned} f_T(\hat{\mathbf{r}})&=(k^2/4\pi\overline{T})\Delta\widetilde{T}_0(\mathbf{K})\cos\,\theta,\\ f_v(\hat{\mathbf{r}})&=(-k^2/2\pi\overline{c})\hat{\mathbf{k}}\cdot\widetilde{\mathbf{v}}_0(\mathbf{K})\cos\,\theta \end{aligned} \tag{2}$$

are the (temperature, velocity) amplitudes, respectively, and

$$\begin{aligned} \Delta\widetilde{T}_0(\mathbf{K})&=\int e^{-i\mathbf{K}\cdot\mathbf{r}'}\Delta T_0(\mathbf{r}')d^3r',\\ \widetilde{\mathbf{v}}_0(\mathbf{K})&=\int e^{-i\mathbf{K}\cdot\mathbf{r}'}\mathbf{v}_0(\mathbf{r}')d^3r' \end{aligned} \tag{3}$$

are the Fourier transforms of the turbulent (temperature variation, velocity) fields $(\Delta T_0,\mathbf{v}_0)$, respectively, with the spectral variable $\mathbf{K}$ (dimension, inverse length). In Eqs. (2) and (3), $\mathbf{r}'$ is an integration variable, and

$$\mathbf{K}=k(\hat{\mathbf{r}}-\hat{\mathbf{k}})\rightarrow K=2k\,\sin(\theta/2). \tag{4}$$

Here, $(\hat{\mathbf{r}},\hat{\mathbf{k}})$ are the (observation, incident) directions, respectively, θ is the scattering angle, with $\hat{\mathbf{k}}\cdot\hat{\mathbf{r}}=\cos\,\theta$, and

$$k=2\pi/\lambda=\omega/\overline{c},\quad \overline{c}=(\gamma R_0\overline{T}/M)^{1/2}, \tag{5}$$

where $\overline{T}$ is the average temperature and $\overline{c}$ is the corresponding adiabatic sound speed, with R_0 the universal gas constant, M the molar mass of the atmosphere, and γ the specific heat ratio. Also

$$\Delta T_0(\mathbf{r})\equiv\overline{T}-T_0(\mathbf{r})\rightarrow\int\Delta T_0(\mathbf{r})d^3r=0. \tag{6}$$

As usual, we assume solenoidal turbulent flow,

$$\boldsymbol{\nabla}\cdot\mathbf{v}_0=0\rightarrow\mathbf{v}_0=\boldsymbol{\nabla}\times\mathbf{A}_0, \tag{7}$$

where $\mathbf{A}_0$ is a vector potential field with dimension (length)^2/time. Using this in Eqs. (2) and (3), integrating by parts, and discarding the integral over the surface at infinity yields

$$f_v(\hat{\mathbf{r}})=(-ik^3/2\pi\overline{c})\cos\,\theta(\hat{\mathbf{k}}\times\hat{\mathbf{r}})\cdot\widetilde{\mathbf{A}}_0(\mathbf{K}), \tag{8}$$

where

$$\widetilde{\mathbf{A}}_0(\mathbf{K})\equiv\int e^{-i\mathbf{K}\cdot\mathbf{r}'}\mathbf{A}_0(\mathbf{r}')d^3r'\rightarrow\widetilde{v}_0(\mathbf{K})=i\mathbf{K}\times\widetilde{\mathbf{A}}_0(\mathbf{K}). \tag{9}$$

Equations (2), (6), and (8) reveal that the conventional Born approximation predicts no scattering at $\theta=90°$, no forward temperature scattering, and no forward or backward velocity scattering for solenoidal turbulent velocities, independently of the morphology of the turbulence.[11]

B. Cross sections and spectra

For isotropic homogeneous stationary stochastic turbulence with solenoidal velocity, the Born approximation ensemble average differential scattering cross sections (dimension, area) are given by[12]

$$\overline{\sigma}(\hat{\mathbf{r}})=\overline{\sigma_T}(\hat{\mathbf{r}})+\overline{\sigma_v}(\hat{\mathbf{r}}), \tag{10}$$

where

$$\overline{\sigma_T}(\hat{\mathbf{r}})\equiv\langle|f_T(\hat{\mathbf{r}})|^2\rangle=(k^2/4\pi\overline{T})^2\,\cos^2\,\theta\Phi^T(\mathbf{K}), \tag{11}$$

$$\overline{\sigma_v}(\hat{\mathbf{r}})\equiv\langle|f_v(\hat{\mathbf{r}})|^2\rangle=(k^2/2\pi\overline{c})^2\,\cos^2\,\theta\hat{k}_i\hat{k}_j\Phi^v_{ij}(\mathbf{K}). \tag{12}$$

Here

$$\Phi^T(\mathbf{K})\equiv\langle|\Delta\widetilde{T}_0(\mathbf{K})|^2\rangle, \tag{13}$$

$$\Phi^v_{ij}(\mathbf{K})\equiv\langle\widetilde{v}_{0i}(\mathbf{K})\widetilde{v}^*_{0j}(\mathbf{K})\rangle=(\delta_{ij}-\hat{K}_i\hat{K}_j)F(K) \tag{14}$$

are called the (temperature, velocity) spectra, respectively; (Φ^T,F) are functions of K. The last equality in Eq. (14) follows from Eq. (9) and the isotropy requirement that $\langle\widetilde{A}_{0p}(\mathbf{K})\widetilde{A}^*_{0q}(\mathbf{K})\rangle$ must equal δ_{pq} times a function of K for an otherwise unrestricted vector field.

The standard treatment due to Kolmogorov finds by dimensional analysis that temperature and velocity structure functions should be proportional to $r^{2/3}$ over some range of r called the inertial range, where r is the distance between

any two locations in space.[1] Assuming (incorrectly of course) that this proportionality is valid for all r, the following expressions, called the Kolmogorov spectra, are easily obtained:

$$\Phi^T(K)=8.19V_T C_T^2 K^{-11/3}, \tag{15}$$

$$\Phi_{ij}^v(K)=(\delta_{ij}-\hat{K}_i\hat{K}_j)(15.1V_T C_v^2 K^{-11/3}), \tag{16}$$

where the coefficients (C_T^2, C_v^2) are called the structure parameters, and V_T is the volume containing turbulence.[13]

II. TURBULE MODELS

We define a stationary turbule as a localized eddy or vortex, characterized by a location $\mathbf{b}$, a scale length a, and flow velocity and temperature variation fields $(\mathbf{v}_0, \Delta T_0)$ that go to zero very rapidly for $|\mathbf{r}-\mathbf{b}|>a$. We treat the location and the morphological parameters in the fields as stochastic variables. We also assume that the flow velocity for each turbule is solenoidal, as in Eq. (7). We take the functional forms of $(\mathbf{A}_0, \Delta T_0)$ for a single turbule as

$$\mathbf{A}_0(\mathbf{r})=a\mathbf{A}(\boldsymbol{\xi}), \quad \Delta T_0(\mathbf{r})=T(\boldsymbol{\xi}), \quad \boldsymbol{\xi}\equiv(\mathbf{r}-\mathbf{b})/a. \tag{17}$$

These are general scalable forms, chosen for convenience so that the functions $\mathbf{A}(\boldsymbol{\xi}), T(\boldsymbol{\xi})$ have the dimension (velocity, temperature), respectively; they may have arbitrary morphology, as long as they go to zero rapidly for $\xi\equiv|\boldsymbol{\xi}|>1$.

We shall need expressions for the scattering amplitudes of an individual turbule in terms of the scaled functions $[\mathbf{A}(\boldsymbol{\xi}), T(\boldsymbol{\xi})]$. A little algebra, starting with Eqs. (2), (3), and (8), yields the following expressions for the scattering amplitudes of a turbule localized around $\mathbf{r}=\mathbf{b}$:

$$f_{T,v}(\hat{\mathbf{r}},\mathbf{b})=e^{-i\mathbf{K}\cdot\mathbf{b}}f_{T,v}(\hat{\mathbf{r}}), \tag{18}$$

where

$$f_T(\hat{\mathbf{r}})=(k^2a^3/4\pi\overline{T})\cos\,\theta\widetilde{T}(\mathbf{K}a), \tag{19}$$

$$f_v(\hat{\mathbf{r}})=-i(k^3a^4/2\pi\overline{c})\cos\,\theta(\hat{\mathbf{k}}\times\hat{\mathbf{r}})\cdot\widetilde{\mathbf{A}}(\mathbf{K}a) \tag{20}$$

are the scattering amplitudes for the same turbule with $\mathbf{b}=\mathbf{0}$. Here the scaled Fourier transforms are defined by

$$\widetilde{T}(\mathbf{y})=\int T(\boldsymbol{\xi})e^{i\boldsymbol{\xi}\cdot\mathbf{y}}\,d^3\xi, \quad \widetilde{\mathbf{A}}(\mathbf{y})=\int \mathbf{A}(\boldsymbol{\xi})e^{i\boldsymbol{\xi}\cdot\mathbf{y}}\,d^3\xi. \tag{21}$$

These have dimensions (temperature, velocity), respectively; the variable $\mathbf{y}$ is an arbitrary dimensionless, real vector-valued variable that will be set equal to $\mathbf{K}a$ in applications.

III. SCATTERING BY INDIVIDUAL TURBULES

A. Isotropic ensembles

1. General

It is clear that for a turbule of arbitrary morphology, the average of its temperature field over random orientations is spherically symmetric, as is the average of $\widetilde{T}(\mathbf{y})$ and $|\widetilde{T}(\mathbf{y})|^2$. Using Eqs. (13), (17), and (21), we may thus write the temperature spectrum for an isotropic ensemble turbule as

$$\Phi^T(K)=a^6\langle|\widetilde{T}(\mathbf{K}a)|^2\rangle=\pi^3a^6(\delta T)^2\widetilde{B}_T^2(Ka), \tag{22}$$

where $\widetilde{B}_T(y)$ is a dimensionless function of $y\equiv|\mathbf{y}|$, and δT is a temperature variation amplitude. The factor π^3 is included for convenience.

For an isotropic ensemble, each component of an otherwise unrestricted stochastic vector field must be statistically independent with zero mean, and must obey the same statistics. Thus we may put

$$\begin{aligned}&\langle\widetilde{A}_i(y)\rangle=0,\\ &\langle\widetilde{A}_i(y)\widetilde{A}_j^*(y)\rangle=\tfrac{1}{3}\pi^3\delta_{ij}v^2\widetilde{B}_v^2(y),\end{aligned} \tag{23}$$

where $\widetilde{B}_v(y)$ is a dimensionless function of y, v is a velocity amplitude, and $\pi^3/3$ is introduced for convenience. It is important to note that Eq. (23) yields the usual form for the turbule velocity spectrum as defined by Eq. (14) for isotropic homogeneous turbulence with solenoidal velocities:

$$\Phi_{ij}^v(\mathbf{K})=(\delta_{ij}-\hat{K}_i\hat{K}_j)(\pi^3a^6v^2/3)(Ka)^2\widetilde{B}_v^2(Ka) \tag{24}$$

that is, a projection off $\hat{\mathbf{K}}$ times a function of K. We shall call the functions $(\widetilde{B}_T(y), \widetilde{B}_v(y))$ ''envelope functions.'' They will be localized functions of y, just as the scaled turbules of Eqs. (17) are localized functions of $\boldsymbol{\xi}$.

The ensemble average differential scattering efficiencies for the isotropic ensemble follow from the definition

$$\overline{Q}(\hat{\mathbf{r}})\equiv\overline{\sigma}(\hat{\mathbf{r}})/\pi a^2=\overline{Q_T}(\hat{\mathbf{r}})+\overline{Q_v}(\hat{\mathbf{r}}), \tag{25}$$

where Eqs. (4), (10)–(12), (22), and (24) yield

$$\overline{Q_T}(\hat{\mathbf{r}})\equiv\overline{\sigma_T}(\hat{\mathbf{r}})/\pi a^2=(\delta T/4\overline{T})^2(ka)^4\,\cos^2\,\theta\widetilde{B}_T^2(Ka), \tag{26}$$

$$\begin{aligned}\overline{Q_v}(\hat{\mathbf{r}})&\equiv\overline{\sigma_v}(\hat{\mathbf{r}})/\pi a^2\\ &=\tfrac{1}{3}(v/2\overline{c})^2(ka)^6\,\cos^2\,\theta\,\sin^2\,\theta\widetilde{B}_v^2(Ka).\end{aligned} \tag{27}$$

These expressions exhibit no dependence on the azimuthal scattering angle φ, as expected for isotropic ensemble averages. They also reveal that the member turbules of an isotropic ensemble, which individually may have arbitrary morphology, may be replaced by member turbules of very high symmetry. First, each member turbule could be characterized by a spherically symmetric temperature field $T(\xi)$, which by Eq. (21) yields a spherically symmetric and real-valued $\widetilde{T}(y)$ and thus determines $\widetilde{B}_T(y)$ according to Eq. (22). So if $\widetilde{B}_T(y)$ is chosen, then $T(\xi)$ is determined: We need only one such member turbule for a given scale length. Second, for the velocity, consider the choice

$$\mathbf{A}(\boldsymbol{\xi})=a\boldsymbol{\Omega}f(\xi)\rightarrow\widetilde{\mathbf{A}}(\mathbf{y})=a\boldsymbol{\Omega}\widetilde{f}(y), \tag{28}$$

with $\boldsymbol{\Omega}$ a constant vector (dimension, inverse time) having random direction in 4π; $\widetilde{f}(y)$ is real valued. Then

$$\langle\Omega_i\Omega_j\rangle=\delta_{ij}\Omega^2/3, \quad \langle\Omega_i\rangle=0 \tag{29}$$

and, if we choose

$$(\widetilde{f}(y))^2=\pi^3\widetilde{B}_v^2(y), \quad \Omega a=v, \tag{30}$$

then we get Eq. (23) immediately. Now, consider the physical significance of the choice (28). With Eqs. (7) and (17) this yields

$$\mathbf{v}_0=\boldsymbol{\Omega}\times(\mathbf{r}-\mathbf{b})(-\xi^{-1}f'(\xi)), \tag{31}$$

a velocity field corresponding to nonuniform rotation of a spherically symmetric structure about an axis along $\hat{\boldsymbol{\Omega}}$ through the point **b**. Clearly, $\boldsymbol{\Omega}$ corresponds to a randomly oriented angular velocity of magnitude v/a.

Therefore we may conclude that any isotropic ensemble of one-turbule systems whose members have arbitrary morphology may be replaced by an ensemble of spherically symmetric nonuniformly rotating one-turbule systems with randomly directed rotation axes. That is, the replacement ensemble produces the same temperature and velocity spectra as the original ensemble, and thus the same ensemble average scattering efficiencies.

2. Comparable turbules

The only freedom left in isotropic ensembles is the choice of the envelope functions and the amplitudes $(\delta T,v)$. We need to be able to compare the scattering produced by isotropic ensembles of the same scale length but different envelope functions. Since the amplitudes are adjustable, we may normalize (B_T,B_v) as desired; for convenience, we choose[14]

$$\tilde{B}_T(0)=\tilde{B}_v(0)=1. \tag{32}$$

We define comparable turbules to be those having the same ''rms radius'' $R_{\rm rms}$. We define this by writing

$$R_{\rm rms}^2=a^2\int \xi^2B(\xi)d^3\xi \Big/ \int B(\xi)d^3\xi, \tag{33}$$

where $B(\xi)$ is the inverse Fourier transform of $\tilde{B}(y)$. We will also take $\tilde{B}_v(y)=\tilde{B}_T(y)$ for convenience, which makes $R_{\rm rms}$ well-defined.

In Sec. IV, we shall show that for isotropic ensembles having different envelope functions, with each ensemble member containing many self-similar turbules of many different scale lengths, the amplitudes may always be adjusted to produce the same scattering over some range of scattering angles.

3. Examples

a. Gaussian envelope. For convenience, we let

$$\tilde{B}_T(y)=\tilde{B}_v(y)\equiv\tilde{B}(y)=e^{-y^2/4}. \tag{34}$$

Such forms have been used in the literature.[8,9] Equation (33) then yields

$$R_{\rm rms}/a=\sqrt{3/2}. \tag{35}$$

The Born approximation ensemble average differential scattering efficiencies follow from Eqs. (26) and (27):

$$\overline{Q_T}(\hat{\mathbf{r}})=(\delta T/4\overline{T})^2(ka)^4\cos^2\theta \times\exp[-(ka)^2(1-\cos\theta)], \tag{36}$$

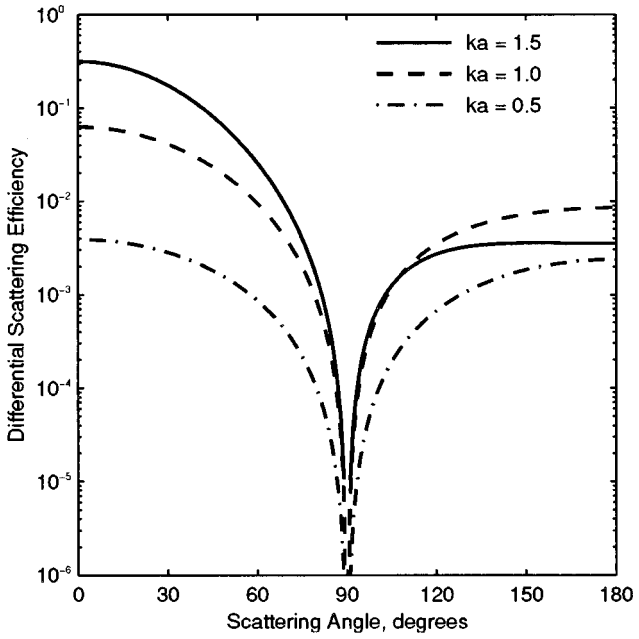


FIG. 1. Differential scattering efficiencies for an isotropic Gaussian temperature turbule ensemble.

$$\overline{Q_v}(\hat{\mathbf{r}})=\tfrac{1}{3}(v/2\overline{c})^2(ka)^6\sin^2\theta\cos^2\theta \times\exp[-(ka)^2(1-\cos\theta)]. \tag{37}$$

These functions are plotted in Figs. 1 and 2 as functions of scattering angle θ for several different size parameters (ka), normalized to unit values of $\delta T/T$ and $v/\overline{c}$. These expressions yield exactly the same dependence of the scattering on θ, k, and a as the expressions in Ref. 7 for a Gaussian spectrum of temperature and wind velocity fluctua-

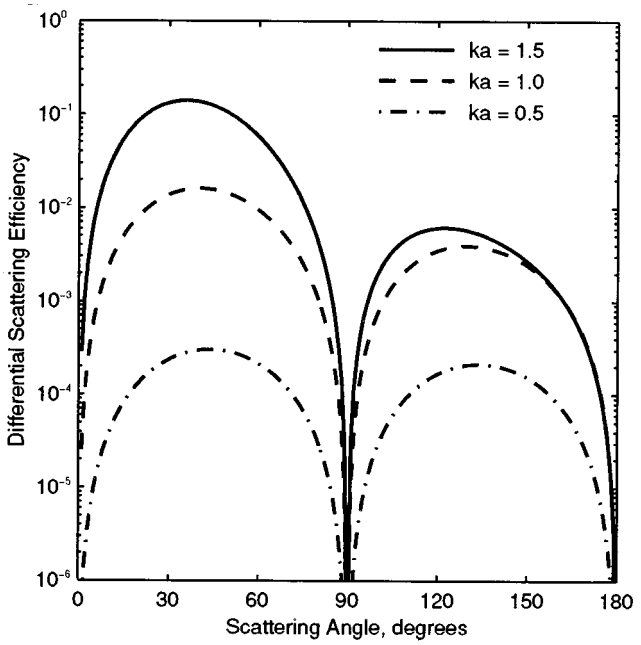


FIG. 2. Differential scattering efficiencies for an isotropic Gaussian velocity turbule ensemble.

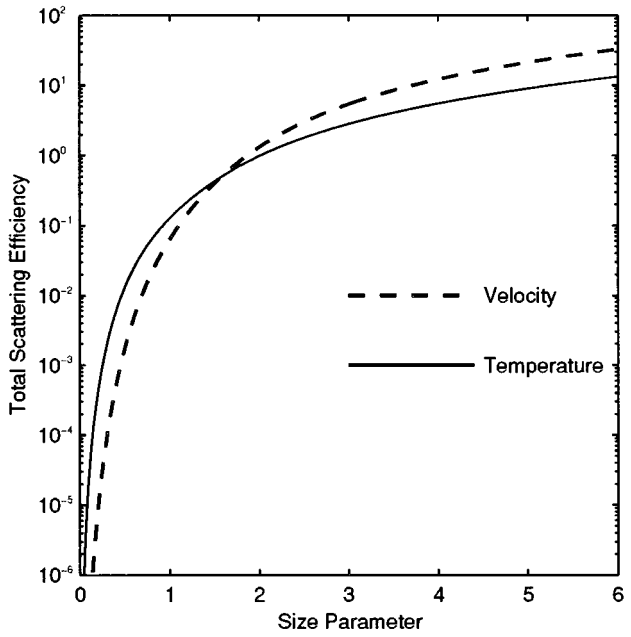


FIG. 3. Total scattering efficiencies for isotropic Gaussian turbule ensembles as functions of size parameter ka.

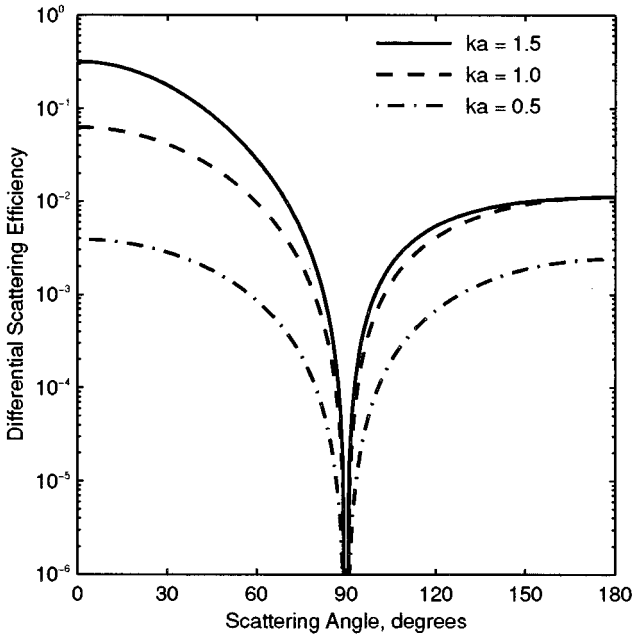


FIG. 4. Differential scattering efficiencies for an isotropic exponential temperature turbule ensemble.

tions. That is, a Gaussian envelope function for an isotropic ensemble of one-turbule systems of a single scale length is the same as a Gaussian spectrum for isotropic turbulence of a single scale length, as would be expected.

The total average efficiencies are given by the integrals of Eqs. (36) and (37) over 4π solid angle. A little algebra yields

$$\overline{Q_T}=2\pi(\delta T/4\overline{T})^2[(x-2+2/x)-e^{-2x}(x+2+2/x)], \tag{38}$$

$$\overline{Q_v}=(4\pi/3)(v/2\overline{c})^2x[(1+e^{-2x})(1+12/x^2) -(1-e^{-2x})(5/x+12/x^3)], \tag{39}$$

where $x=(ka)^2$.

In Fig. 3, these functions are plotted versus ka for unit values of $\delta T/\overline{T}$ and $v/\overline{c}$. For small ka, these efficiencies reduce to

$$\overline{Q_T}\approx(4\pi/3)(\delta T/4\overline{T})^2(ka)^4, \qquad \overline{Q_v}\approx\tfrac{1}{3}(8\pi/15)(v/2\overline{c})^2(ka)^6. \tag{40}$$

The quite different dependence on ka is striking: If $(\delta T/\overline{T}) \approx(v/\overline{c})$ then $\langle Q_v\rangle$ is much smaller than $\langle Q_T\rangle$ for small ka. For very large ka, both efficiencies are proportional to $(ka)^2$, as can be seen easily in Eqs. (38) and (39). As is well-known, the Born approximation fails in the limit of arbitrarily large ka; the actual total scattering efficiencies must be independent of ka in this limit. This means that the Born approximation should not be used to describe the scattering by very large scale turbules.

b. Exponential envelope. We let

$$\widetilde{B}_T(y)=\widetilde{B}_v(y)\equiv\widetilde{B}(y)=(1+y^2/\alpha^2)^{-3}, \tag{41}$$

where α is a parameter to be adjusted. This corresponds to

$$B(\xi)\equiv(2\pi)^{-3}\int e^{i\boldsymbol{\xi}\cdot\mathbf{y}}\widetilde{B}(y)d^3y =(\alpha^3/32\pi)(1+\alpha\xi)e^{-\alpha\xi}. \tag{42}$$

The factor $(1+\alpha\xi)$ is necessary to ensure that the velocity field is bounded at $\xi=0$. Equations (33) and (35) yield

$$R_{\rm rms}/a=\sqrt{18}/\alpha\rightarrow\alpha=\sqrt{12} \tag{43}$$

in order that the exponential and Gaussian model turbules have the same effective size.

The ensemble average differential scattering efficiencies then follow from Eqs. (26) and (27):

$$\overline{Q_T}(\hat{\mathbf{r}})=(\delta T/4\overline{T})^2(ka)^4 \times\cos^2\theta[1+\tfrac{1}{6}(ka)^2(1-\cos\theta)]^{-6}, \tag{44}$$

$$\overline{Q_v}(\hat{\mathbf{r}})=\tfrac{1}{3}(v/2c)^2(ka)^6\sin^2\theta \times\cos^2\theta[1+\tfrac{1}{6}(ka)^2(1-\cos\theta)]^{-6}. \tag{45}$$

Note that these are the same as the Gaussian results (36) and (37) for small ka; our normalization $\widetilde{B}(0)=1$ ensured this. These efficiencies are plotted in Figs. 4 and 5 as functions of θ for several size parameters ka, for unit values of $\delta T/\overline{T}$ and $v/\overline{c}$.

Analytic results for the total efficiencies versus ka are quite lengthy, so we do not quote them. The limits for small ka are of course the same as the results (40) for the Gaussian envelope. The exponential model curves in Fig. 6 were obtained from analytic forms of integrals of Eqs. (44) and (45) as found by Mathematica.[15] Just as for the Gaussian (or any other model), for very large ka, the total efficiencies are proportional to $(ka)^2$; this is a generic defect of the Born approximation.

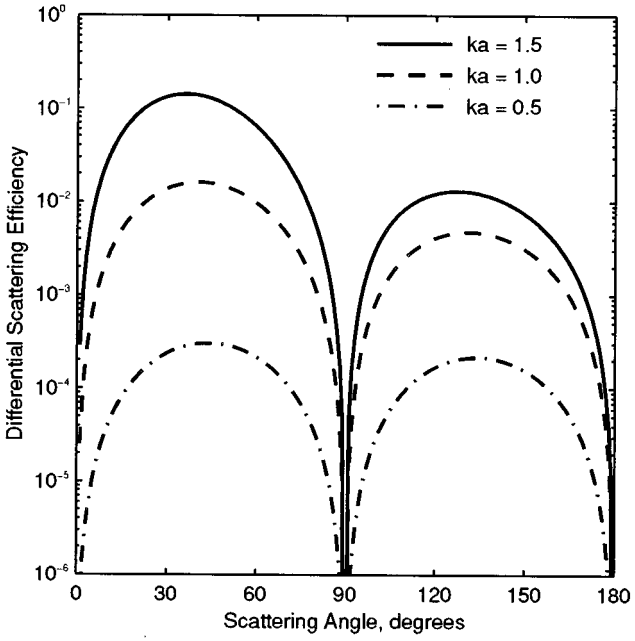


FIG. 5. Differential scattering efficiencies for an isotropic exponential velocity turbule ensemble.

B. Individual turbule model

We include this in order to illustrate the azimuthal dependence of the velocity scattering from even the simplest model turbule having a solenoidal velocity field, which is of the type given by Eq. (28). We write $\hat{\mathbf{\Omega}}$ in terms of its polar and azimuthal angles $(\theta_\Omega, \varphi_\Omega)$,

$$\hat{\mathbf{\Omega}} = \mathbf{e}_1 \sin\theta_\Omega \cos\varphi_\Omega + \mathbf{e}_2 \sin\theta_\Omega \sin\varphi_\Omega + \mathbf{e}_3 \cos\theta_\Omega,$$

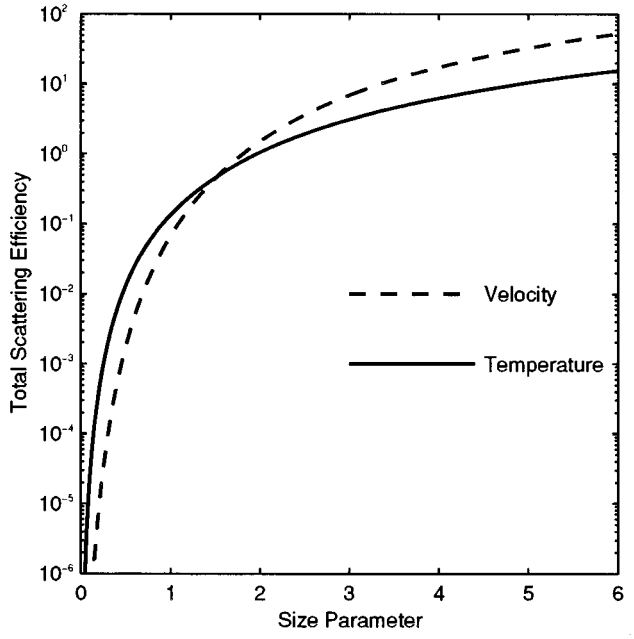


FIG. 6. Total scattering efficiencies for isotropic exponential turbule ensembles as functions of size parameter ka.

where the $\mathbf{e}_i$ are the Cartesian basis vectors, and we choose $\mathbf{k}$ in the z direction, and use

$$\hat{\mathbf{r}} = \mathbf{e}_1 \sin\theta\cos\varphi + \mathbf{e}_2 \sin\theta\sin\varphi + \mathbf{e}_3 \cos\theta,$$

plus Eqs. (21) and (20) and $Q_v = |f_\nu|^2/\pi a^2$, to get

$$Q_v(\hat{\mathbf{r}}) = (ka)^6 (v/2c)^2 \sin^2\theta \cos^2\theta \sin^2\theta_\Omega \sin^2(\varphi - \varphi_\Omega). \tag{46}$$

This is not only zero at $\theta = (0, \pi/2, \pi)$, but also if $\theta_\Omega = (0, \pi)$, that is, when Ω is along the $\hat{\mathbf{k}}$ direction, and also if $\varphi = \varphi_\Omega \pm \pi$. This strong azimuthal dependence already results from the simplest possible turbule with solenoidal velocity. If we average over random directions of $\hat{\mathbf{\Omega}}$, using

$$(4\pi)^{-1} \int_0^{2\pi} d\varphi_\Omega \int_0^{\pi} d\theta_\Omega \sin\theta_\Omega [\sin^2\theta_\Omega \sin^2(\varphi_\Omega - \varphi)]$$

$$= \tfrac{1}{3}$$

of course we get Eq. (27) for the isotropic ensemble.

IV. SCATTERING BY ISOTROPIC HOMOGENEOUS ENSEMBLES

A. General

Here we model isotropic homogeneous turbulence contained in a volume V_T as an isotropic homogeneous ensemble of many-turbule systems. Each system contains self-similar stationary turbules of arbitrary morphology and many different scale lengths, located randomly but uniformly in V_T. On the average, in each system we allow N_α turbules of scale length a_α, $\alpha = (1, N)$, where N is the total number of different scale lengths in the system, so that $N_T = \Sigma_\alpha N_\alpha$ is the ensemble average number of turbules per system in volume V_T. We regard N_α and N as fixed in this paper. We define a_1 as the largest scale length in the system, and a_N as the smallest; these lengths define the outer and inner scales of the turbulence, usually designated as L_0 and l_0, respectively. Each turbule is represented by the general localized scalable fields of Eqs. (17). The total turbulent fields in each system are each just superpositions of the fields of the N_T constituent turbules.

For our purposes we do not need to provide a detailed derivation of ensemble averages. Instead, we simply use the known result for scattering by collections of randomly placed particles: The total scattering is incoherent, except near the forward direction. Therefore, the average total scattering cross section is just the sum

$$\overline{\sigma}(\hat{\mathbf{r}}) = \sum_{n=1}^{N_T} \overline{\sigma_n}(\hat{\mathbf{r}}) = \sum_{\alpha=1}^{N} N_\alpha \overline{\sigma_\alpha}(\hat{\mathbf{r}}) \tag{47}$$

except for a narrow cone near the forward direction; the larger V_T, the smaller that cone. The second equality follows because the isotropic ensemble average cross section for each turbule of size α is the same. This equation also applies separately for the temperature and velocity cross sections. We will not calculate forward or very near forward scattering, which we already know approaches zero.

It follows immediately that the spectra also sum incoherently, except near $K=0$:

$$\Phi^T(K)=\sum_{\alpha=1}^{N} N_\alpha \Phi_\alpha^T(K), \quad \Phi_{ij}^v(\mathbf{K})=\sum_{\alpha=1}^{N} N_\alpha \Phi_{\alpha ij}^v(\mathbf{K}). \tag{48}$$

Using Eqs. (22) and (24) for the isotropic ensemble of one-turbule systems, we obtain

$$\Phi^T(K)=\pi^3 \sum_{\alpha=1}^{N} N_\alpha \alpha_\alpha^6 (\delta T_\alpha)^2 \tilde{B}_T^{\,2}(Ka_\alpha), \tag{49}$$

$$\Phi_{ij}^v(\mathbf{K})=(\delta_{ij}-\hat{K}_i\hat{K}_j) \times(\pi^3/3)\sum_{\alpha=1}^{N} N_\alpha a_\alpha^6 v_\alpha^2 (Ka_\alpha)^2 \tilde{B}_v^{\,2}(Ka_\alpha). \tag{50}$$

The ensemble average scattering cross sections for θ not near 0° are then obtained by using these in Eqs. (11) and (12).

B. Scaling

In order to describe the complete ensemble, we assume that the quantities $(N_\alpha, \Delta T_\alpha, v_\alpha)$ scale with a_α. We put

$$\frac{N_\alpha}{N_1}=\left(\frac{a_\alpha}{a_1}\right)^{-\beta}, \quad \left(\frac{\delta T_\alpha}{\delta T_1}\right)=\left(\frac{a_\alpha}{a_1}\right)^{\gamma}, \quad \left(\frac{v_\alpha}{v_1}\right)=\left(\frac{a_\alpha}{a_1}\right)^{\nu}, \tag{51}$$

where (β,γ,ν) are parameters to be determined. In addition, we must decide how to relate the scale lengths to the index α. One relation that has been used is[16]

$$a_\alpha=a_1 e^{-\mu(\alpha-1)}, \quad \mu>0, \tag{52}$$

where μ is a parameter that is determined by N, the number of scale lengths, and the ratio (a_N/a_1) of inner to outer scale lengths:

$$\mu=-(N-1)^{-1}\ln m, \quad m\equiv a_N/a_1. \tag{53}$$

The relation (52) implies that the scale lengths from a geometric sequence, in which

$$a_2/a_1=a_3/a_2=\cdots=e^{-\mu}. \tag{54}$$

This is a kind of fractal scaling.

A general power-law scaling relation is given by

$$a_\alpha=a_1(1+q\mu(\alpha-1))^{-1/q}, \quad q>0, \quad \mu>0, \tag{55}$$

where here μ is determined in terms of (m,q) by

$$\mu=q^{-1}(m^{-q}-1)/(N-1). \tag{56}$$

Equations (55) and (56) actually reduce to Eqs. (52) and (53), respectively, when $q\to 0$, so in what follows we may use just Eqs. (55) and (56) with $q\geqslant 0$. We will also make use of the Kolmogorov concept of energy transfer in fully developed turbulence,[1,2] in which the kinetic energy transfer rate (dimension energy/time) from eddies of scale length a_α to the next smaller is the same for all sizes $1<\alpha<N-1$. By dimensional analysis and from the fluid equations, this rate is

$$\dot{\mathscr{E}}_{K\alpha}=(C)(N_\alpha)(a_\alpha^3)(v_\alpha^2)(v_\alpha/a_\alpha), \tag{57}$$

where C is a constant. That is, the rate is proportional to the number of turbules of size a_α in V_T, the volume of each, the kinetic energy per unit mass $v_\alpha^2/2$ of each, and the characteristic rate of transfer v_α/a_α.

Requiring $\dot{\mathscr{E}}_{K\alpha}$ to be independent of a_α and using Eq. (51) yields

$$\beta=3\nu+2. \tag{58}$$

In the atmosphere, the ratios $|\Delta T_0|/\overline{T}$ and $v_0/\overline{c}$ are usually of the same order.[1] This implies that our turbule temperature variation amplitudes δT_α should be proportional to v_α, whereby Eq. (51) yields

$$\gamma=\nu. \tag{59}$$

C. Spectra and cross sections

1. General forms

We insert the scaling relations (51) into Eqs. (49) and (50), and convert the sums to integrals, using

$$\sum_{\alpha=1}^{N} \approx \int_1^N d\alpha=\int_{a_1}^{a_N}\frac{da}{da/d\alpha}, \quad \frac{da}{d\alpha}=-\mu a_1^{-q} a^{1+q}, \tag{60}$$

where the last equality results from Eq. (55). This yields

$$\Phi^T(K)=(\pi^3 N_1(\delta T_1)^2 a_1^6/\mu)x^{-P_T} J_{P_T-1}^T(mx,x), \tag{61}$$

$$\Phi_{ij}^v(\mathbf{K})=(\delta_{ij}-\hat{K}_i\hat{K}_j) \times(\pi^3 N_1 v_1^2 a_1^6/3\mu)x^{-P_v} J_{P_v+1}^v(mx,x), \tag{62}$$

where (m,μ) are defined by Eqs. (52) and (53) or Eqs. (54) and (55), and the quantities $J^{T,v}$ and (P_T,P_v) are defined by

$$J_s^{T,v}(mx,x)\equiv\int_{mx}^{x} dy\; y^s \tilde{B}_{T,v}^{\,2}(y), \quad x\equiv Ka_1, \tag{63}$$

$$P_T=P_v\equiv P=6+2\nu-\beta-q=4-\nu-q, \tag{64}$$

where Eqs. (58) and (59) were used and s is an index that may be P_T-1 or P_v+1. If we regard K as an independent Fourier transform variable here, then $(x=Ka_1,\ mx=Ka_N)$ are the (outer, inner) scale size parameters of the turbulence, respectively.

It is important to determine qualitatively how the integral of Eq. (63) depends on (s,m,x). In all cases, the envelope functions $\tilde{B}_{T,v}^{\,2}(y)$ are bounded for all y and go to zero rapidly for large y, because individual turbules are localized. Thus for $s\gtrsim 1$ the integrands are peaked functions and are negligibly small for $0<y<y_{s-}$ and $y>y_{s+}$, where the values of $y_{s\pm}$ are dependent on the particular forms of the envelope functions and the index s. It is clear that the integrals are essentially zero for $x\lesssim y_{s-}$ or $mx\gtrsim y_{s+}$, and also that

$$J_s^{T,v}(mx,x)\approx J_s^{T,v}(0,\infty)=\text{const}, \qquad x_{\min}\equiv y_{s+}<x<y_{s-}/m\equiv x_{\max}. \tag{65}$$

Thus $(x_{\min},x_{\max})$ are effective inertial range boundaries on the dimensionless variable $x=Ka_1$. For this range of x, Eqs. (61) and (62) show that the spectra have a common power-

law dependence x^{-P}, with P given by Eq. (64), independently of the detailed form of the envelope functions. For the transition regions $y_{s-} \lesssim x \lesssim y_{s+}$ and $y_{s-}/m \lesssim x \lesssim y_{s+}/m$, the spectra depart markedly from the power law.

2. Kolmogorov spectra

We can match Eqs. (61) and (62) to the Kolmogorov spectra of Eqs. (15) and (16) in the inertial range, by choosing

$$P = \tfrac{11}{3} \rightarrow \nu = \tfrac{1}{3} - q, \quad \beta = 3(1-q), \tag{66}$$

where Eqs. (58) and (64) yield the last equalities, and then identifying the following expressions for the structure parameters:

$$C_T^2 = (3.78/\mu)(N_1 a_1^3/V_T)(\delta T_1/a_1^{1/3})^2 J_{8/3}^T(0,\infty), \tag{67}$$

$$C_v^2 = (0.68/\mu)(N_1 a_1^3/V_T)(v_1/a_1^{1/3})^2 J_{14/3}^v(0,\infty). \tag{68}$$

If the fractal length scaling $q=0$ of Eqs. (52) and (53) is used, then $\nu = 1/3$, the standard result; that is, Eq. (51) shows that $(\delta T_\alpha, v_\alpha) \propto a_\alpha^{1/3}$, and also that $N_\alpha a_\alpha^3 = \text{const}$, i.e., that turbule packing fractions are scale invariant.[17] Thus for $q=0$, the factors $(N_1 a_1^3/V_T, \delta T_1/a_1^{1/3}, v_1/a_1^{1/3})$ are also scale invariant, and therefore so are the expressions for C_T^2 and C_v^2; this is not the case for power-law scaling with $q \neq 0$.

3. Morphology dependence

It is important to examine the effects of changing the turbule model envelope functions $\tilde{B}^2(y)$ that appear in the integrals $J_s(mx,x)$ of Eq. (63).

We may change an envelope function in two nontrivial ways (an amplitude change is trivial; see Sec. III A 2). One way is to replace $\tilde{B}(y)$ by the same functions of a ''stretched'' argument, $\tilde{B}(y) \rightarrow \tilde{B}(\gamma y)$. This is equivalent to changing the spectrum of scale lengths in the ensemble, such that $a_\alpha \rightarrow a'_\alpha = \gamma a_\alpha$, and altering $(\delta T_1, v_1)$ appropriately to keep the same values of $\Phi^T(K), \Phi_{ij}^v(K)$ in the inertial range. But the boundaries of the inertial range will be shifted; that is, Eqs. (63) and (65) yield

$$K'_{\min} = y_{s+}/a'_1 = K_{\min}/\gamma, \quad K'_{\max} = K_{\max}/\gamma. \tag{69}$$

Outside the inertial range, the spectra will be changed. However, γ cannot be chosen much different from unity, because then, for example, the largest turbules would have a much different linear extension a'_1 than the presumed known outer scale length of the turbulence.

Another way is to replace a chosen $\tilde{B}(y)$ by a different functional form $\tilde{B}'(y)$. Clearly this will change the boundaries of the inertial ranges, in general, but inside the inertial ranges, $(\delta T_1, v_1)$ can be altered to preserve the previous values of the spectra. Outside the inertial range, the spectra will be changed. Again, the different functional forms should be chosen to have comparable length scales, for example, the same rms radii for each index α, in order to represent the same known outer and inner scales.

Therefore we may conclude the following: The power-law spectra in the inertial ranges are completely insensitive to all changes in turbule morphology, that is, alterations of the envelope functions $\tilde{B}(y)$. Changes in the $\tilde{B}(y)$ irreducibly influence only the boundaries $(K_{\min}, K_{\max})$ of the inertial range and the behavior of the spectra and the scattering cross sections for K outside the inertial range.

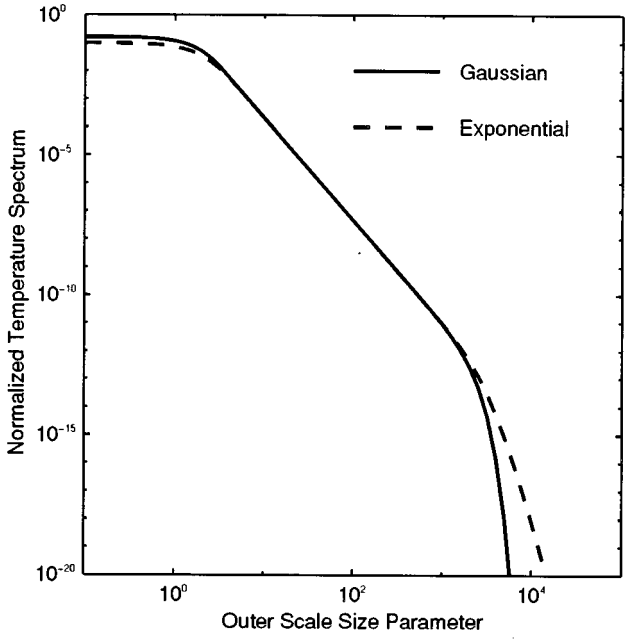


FIG. 7. Normalized isotropic homogeneous ensemble temperature spectra for $m = 10^{-3}$ as functions of outer scale size parameter ka_1.

4. Examples

Equations (34), (41), and (44) considered two example envelope functions, a Gaussian and the Fourier transform of an exponential, that have the same rms turbule radii for each a_α. For the quasi-Kolmogorov spectra, with $P = 11/3$ in Eqs. (61) and (63), we consider the normalized spectra $[F_T(x), F_v(x)]$ defined by

$$\begin{aligned} F_T(x) &\equiv x^{-11/3} J_{8/3}(mx,x)/J_{8/3}(0,\infty), \\ F_v(x) &\equiv x^{-11/3} J_{14/3}(mx,x)/J_{14/3}(0,\infty), \end{aligned} \tag{70}$$

in which we use the same envelope functions $\tilde{B}(y)$ for both temperature and velocity spectra, for convenience. These functions are the factors that determine the boundaries of the inertial ranges of K and the behavior of $\Phi^T(K), \Phi_{ij}^v(K)$ outside the inertial ranges. From Eq. (63), we have for the example envelopes

$$\begin{aligned} J_s^g(mx,x) &= \int_{mx}^{x} dy\ y^s e^{-y^2}, \\ J_s^e(mx,x) &= \int_{mx}^{x} dy\ y^s (1 + y^2/12)^{-6}, \end{aligned} \tag{71}$$

where (g,e) stand for (Gaussian, exponential), respectively. These integrals were evaluated analytically and numerically for $s = (8/3, 14/3)$ and for $m = a_N/a_1 = (10^{-3}, 10^{-4})$, which are realistic values for the ratio of inner to outer scale length.[18]

Figures 7 and 8 are plots of F_T^g and F_T^e vs x for $m = 10^{-3}$ and 10^{-4}, respectively; Figs. 9 and 10 are plots of

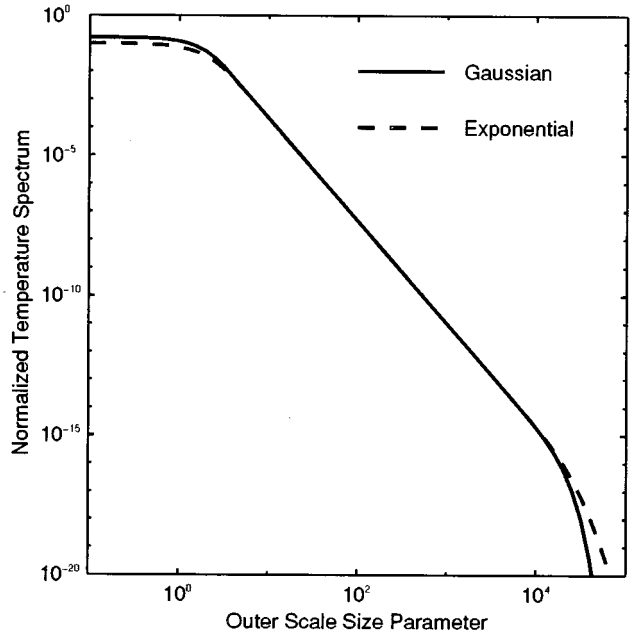


FIG. 8. Normalized isotropic homogenous ensemble temperature spectra for $m=10^{-4}$ as functions of outer scale size parameter ka_1.

F_v^g and F_v^e vs x for $m=10^{-3}$ and 10^{-4}, respectively. The plots of course coincide in the common portions of their inertial ranges, where their slope is $(-11/3)$ on these log-log plots. But the exponential and Gaussian envelopes yield slightly different inertial range boundaries, quite different behavior for $x>x_{\max}$, and the same behavior but different values for $x<x_{\min}$. Here approximate values for the effective inertial range boundaries $x_{\max}$ and $x_{\min}$ for each curve are evident from the plots.

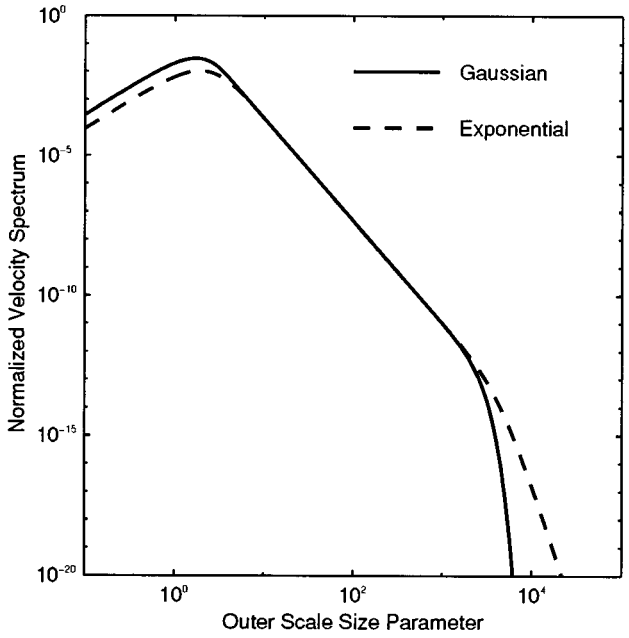


FIG. 9. Normalized isotropic homogeneous ensemble velocity spectra for $m=10^{-3}$ as functions of outer scale size parameter ka_1.

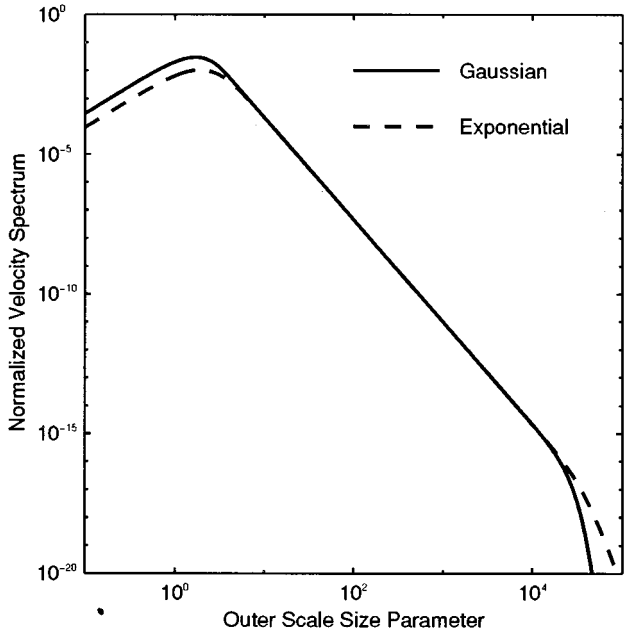


FIG. 10. Normalized isotropic homogeneous ensemble velocity spectra for $m=10^{-4}$ as functions of outer scale size parameter ka_1.

Expressions for scattering cross sections corresponding to these examples are obtained by combining Eqs. (11), (12), (61), (62), (67), (68), and (70). We get

$$Q_T' \equiv \sigma_T / A_v = \cos^2\ \theta F_T(x),$$
$$Q_v' \equiv \sigma_v / A_T = \cos^2\ \theta\ \cos^2(\theta/2) F_v(x), \tag{72}$$

where

$$A_T \equiv 0.052 C_T^2 k^4 V_T a_1^{11/3} / \overline{T^2}, \quad A_v \equiv 0.38 C_v^2 k^4 V_T a_1^{11/3} / \overline{c^2}. \tag{73}$$

Note that (A_T, A_v) have the dimension of area, so the quantities (Q_T', Q_v') are ''scattering efficiencies'' analogous to those defined for individual turbules.

Given values for (m, ka_1), we can use $x = 2ka_1 \sin(\theta/2)$ to calculate inertial range boundaries $(\theta_{\min}, \theta_{\max})$ on θ, putting either or both equal to 180° if $x/2ka_1 \geqslant 1$. Figure 11 is a plot of $\theta_{\min}$ and $\theta_{\max}$ vs ka_1 for $m=10^{-3}$ and $x_{\min}\approx 2.4$, $x_{\max}\approx 900$, the values from Fig. 7 for the Gaussian temperature spectrum ($s=8/3$). Figure 12 is the same for $x_{\min}\approx 2.9$, $x_{\max}\approx 1400$, the values from Fig. 9 for the Gaussian velocity spectrum ($s=14/3$). These figures show that for many ka_1 the inertial range $\theta_{\min} \leqslant \theta \leqslant \theta_{\max}$ is significantly smaller than the full range $0° \leqslant \theta \leqslant 180°$ of scattering angles.

The $(x_{\min}, x_{\max})$ values may also be estimated analytically by solving for $y_{s\pm}$, the values of y at which the integrands of Eq. (71) fall to say e^{-1} of their peak values, and using Eq. (65). Alternatively, we could define $y_{s\pm}$ as the points at which the slopes of the integrands change most rapidly.

In order to exhibit in detail how inertial range boundaries can influence the scattering, in Fig. 13 we plot Q_T' vs θ for the Gaussian envelope functions in Eq. (71), for $m=10^{-3}$, $a_1=10$ m, and two wavelengths, $\lambda_1=10$ m,

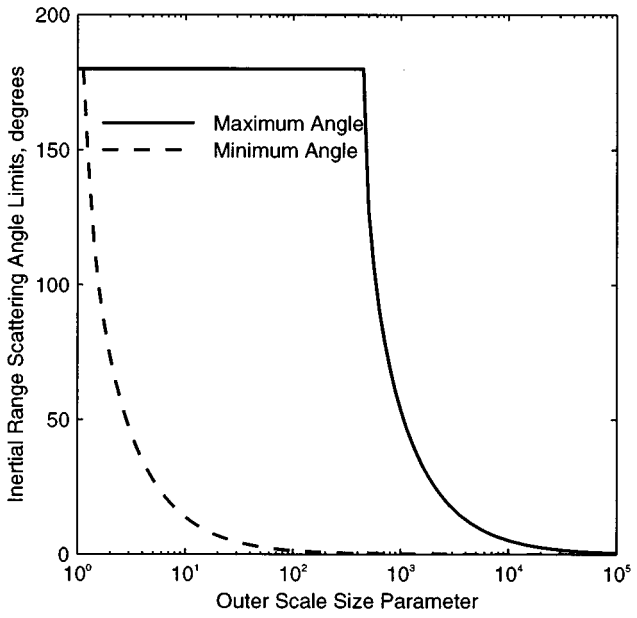


FIG. 11. Inertial range scattering angle boundaries for isotropic homogeneous Gaussian temperature turbule ensembles with $m=10^{-3}$ as functions of outer scale size parameter ka_1.

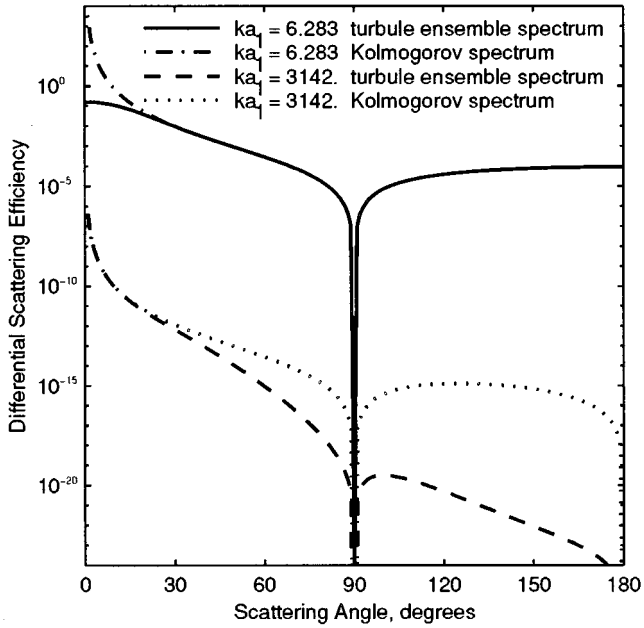


FIG. 13. Differential scattering efficiencies for the Gaussian turbule temperature spectrum compared to those for the Kolmogorov spectrum for $m=10^{-3}$ and different outer scale size parameters ka_1.

$\lambda_2=0.02$ m. In Fig. 14 we do the same for Q_v'. Note that the scattering for θ outside the inertial range is significantly less than would be predicted by the extended inertial range curves (dashed). This is the case in general.

Finally, in order to obtain the temperature structure function $D^T(r)$ from the spectrum $\Phi^T(K)$, we start with the definition

$$D^T(r)\equiv\langle(\Delta T_0(\mathbf{r}_1)-\Delta T_0(\mathbf{r}_2))^2\rangle,\quad r\equiv|\mathbf{r}_1-\mathbf{r}_2|. \tag{74}$$

Using Eqs. (3) and (13), we get easily

$$D^T(r)=(\pi^2V_T)^{-1}\int_0^\infty dK\,K^2\left(1-\frac{\sin Kr}{Kr}\right)\Phi^T(K). \tag{75}$$

Using Eqs. (61), (63), (66), and (67), we then get

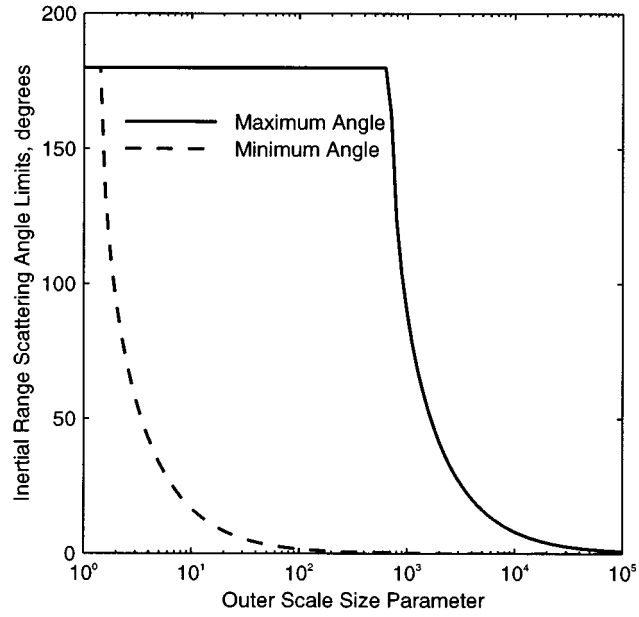


FIG. 12. Inertial range scattering angle boundaries for isotropic homogeneous Gaussian velocity turbule ensembles with $m=10^{-3}$ as functions of outer scale size parameter ka_1.

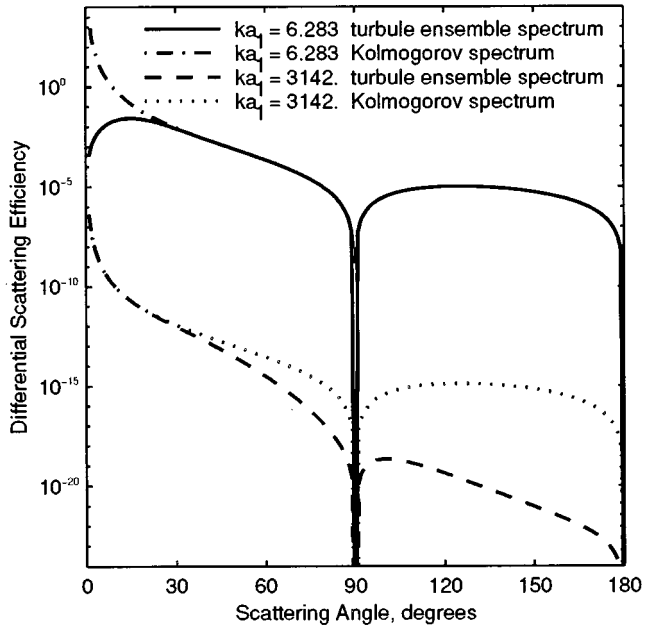


FIG. 14. Differential scattering efficiencies for the Gaussian turbule velocity spectrum compared to those for the Kolmogorov spectrum for $m=10^{-3}$ and different outer scale size parameters ka_1.

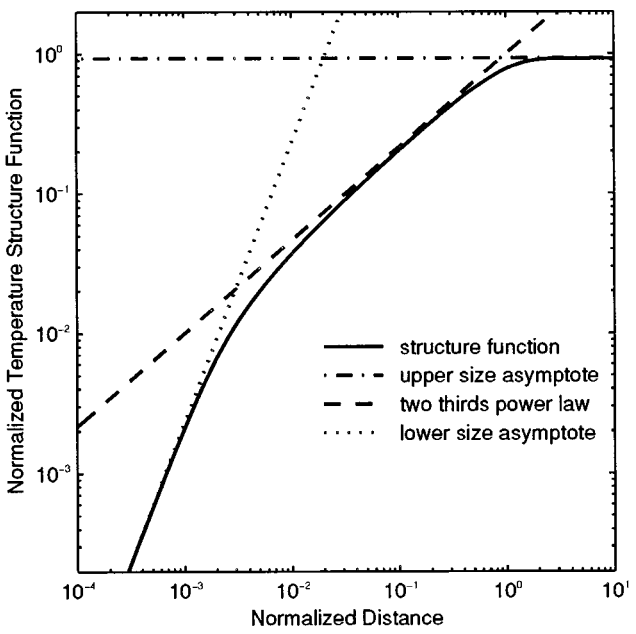


FIG. 15. Normalized structure function for an isotropic homogeneous Gaussian temperature turbule ensemble with $m=10^{-3}$ as a function of normalized distance r/a_1.

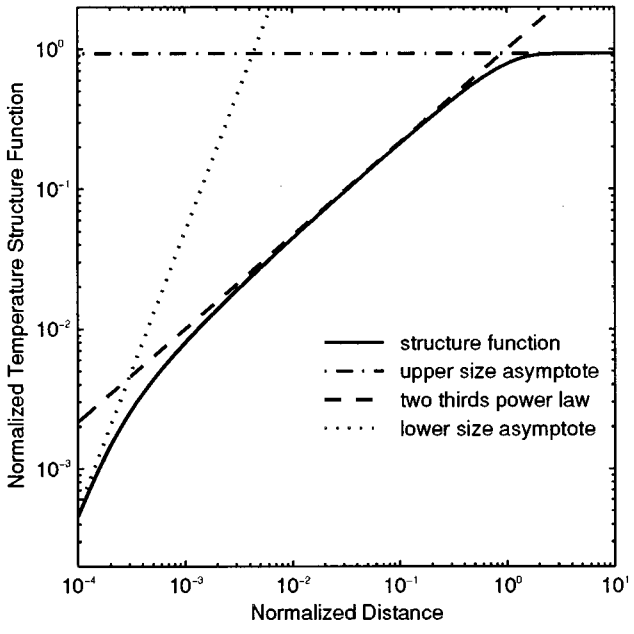


FIG. 16. Normalized structure function for an isotropic homogeneous Gaussian temperature turbule ensemble with $m=10^{-4}$ as a function of normalized distance r/a_1.

$$D^T(r)=a_1^{2/3}C_T^2 I_T(\xi), \quad \xi\equiv r/a_1. \tag{76}$$

Here ξ is a normalized distance in units of a_1, and $I_T(\xi)$ is a dimensionless structure function, defined by

$$I_T(\xi)=(\Lambda J^T_{8/3}(0,\infty))^{-1}\xi^{2/3}\int_0^\infty dz\ z^{-5/3}(1-z^{-1}\sin z) \times J^T_{8/3}(mz/\xi,z/\xi) \tag{77}$$

with

$$\Lambda\equiv\int_0^\infty dz\ z^{-5/3}(1-z^{-1}\sin z)=1.205. \tag{78}$$

From Eq. (76), we see that $I_T(\xi)$ should equal just $\xi^{2/3}$ in the inertial range since $D_T(r)$ should equal $C_T^2 r^{2/3}$ there.

In the Appendix, we show that, for the Gaussian envelope function (34), the expression (77) reduces exactly to

$$I_T(\xi)=0.622\xi^{2/3}\int_\xi^{\xi/m} dy\ y^{-5/3}(1-e^{-y^2/2}). \tag{79}$$

The integral here involves incomplete Gamma functions. In obtaining the log-log plots of Figs. 15 and 16 for $I_T(\xi)$ for $m=(10^{-3},10^{-4})$, respectively, these functions were evaluated using MATLAB.

These plots clearly reveal the inertial ranges and the correct asymptotic behaviors of these dimensionless structure functions. From Eq. (79), their asymptotic limits (for $m\ll 1$) are

$$I_T(\xi)\underset{\xi\to 0}{\sim}0.233m^{-4/3}\xi^2, \quad I_T(\xi)\underset{\xi\to\infty}{\sim}0.932. \tag{80}$$

The inertial range of ξ in Eq. (79) corresponds to those values of ξ for which the limits on the integral may be changed from $(\xi,\xi/m)$ to $(0,\infty)$, respectively, with negligible error. Since the integrand in Eq. (79) is peaked with a maximum at $y_m\approx 0.35$, then for ξ such that $\xi\ll y_m$ and $\xi/m\gg y_m$, this should occur.

As a check, using

$$\int_0^\infty dy\ y^{-5/3}(1-e^{-y^2/2})=\frac{3}{2}\int_0^\infty dy\ y^{1/3}e^{-y^2/2}=1.612, \tag{81}$$

we get from Eqs. (79) and (81) that $I_T(\xi)=(1.00)\xi^{2/3}$ in the inertial range, as expected.

The velocity structure function is defined by

$$D^v_{ij}(\mathbf{r})\equiv\langle(v_{0i}(\mathbf{r}_1)-v_{0i}(\mathbf{r}_2))(v_{0j}(\mathbf{r}_1)-v_{0i}(\mathbf{r}_2))\rangle. \tag{82}$$

Using Eqs. (3), (14), (62), (63), (66), and (68), we get

$$D^v_{ii}=(11/3)a_1^{2/3}C_v^2 I_v(\xi) \tag{83}$$

for the trace of Eq. (82), where

$$I_v(\xi)\equiv(\Lambda J^v_{14/3}(0,\infty))^{-1}\xi^{2/3} \times\int_0^\infty dz\ z^{-5/3}\left(1-\frac{\sin z}{z}\right)J^v_{14/3}(mz/\xi,z/\xi) \tag{84}$$

is a dimensionless velocity structure function. Since $D^v_{ii}(r)$ should equal $(11/3)C_v^2 r^{2/3}$ in the inertial range,[1] $I_v(\xi)$ should equal $\xi^{2/3}$ there.

For the Gaussian envelope function (34), the method used in the Appendix for $I_T(\xi)$ yields

$$I_v(\xi)=0.508\xi^{2/3}\int_\xi^{\xi/m} dy[y^{5/3}(1-e^{-y^2/2})+\tfrac{1}{3}y^{1/3}e^{-y^2/2}]. \tag{85}$$

As a check, putting $\xi\to 0$, $\xi/m\to\infty$ in the limits of the integral for the inertial range and using Eq. (81) yields $I_v(\xi)=1.00\xi^{2/3}$ there, as expected. Figures 17 and 18 are log-log

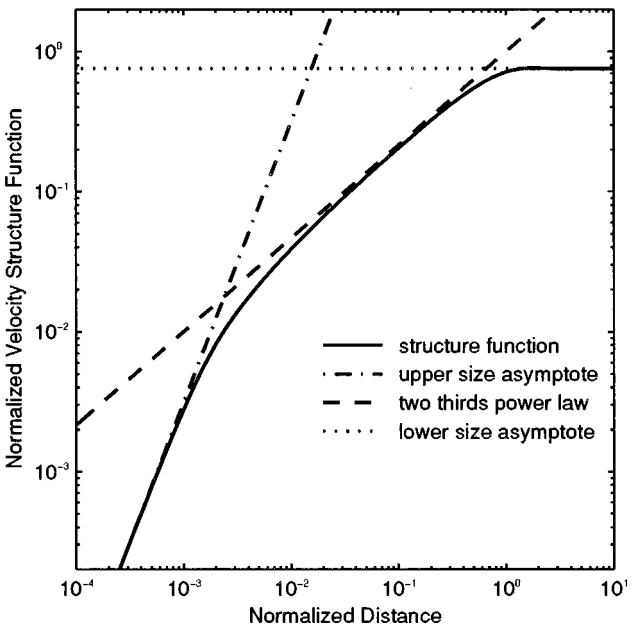


FIG. 17. Normalized structure function for an isotropic homogeneous Gaussian velocity turbule ensemble with $m=10^{-3}$ as a function of normalized distance r/a_1.

plots of Eq. (85) for $m=(10^{-3},10^{-4})$, respectively. Just as for the temperature structure functions, the log-log plots clearly reveal the extent of the inertial ranges and the correct asymptotic behaviors.

V. SUMMARY AND DISCUSSION

In this paper, we have modeled isotropic homogeneous turbulence as a superposition of randomly located randomly oriented self-similar turbules of many different sizes, in which number densities and temperature and velocity amplitudes of the turbules scale with size, and sizes scale according to a power law in the size index. We derived expressions for the ensemble average acoustic scattering amplitudes and cross sections for several turbule ensembles.

The principal results are in Sec. IV, where we showed that

(i) The temperature and velocity spectra of the turbulence and the associated acoustic scattering cross sections have a power-law dependence for some range of the spectral variable K.

(ii) The boundaries of this so-called inertial range are slightly dependent on the model turbule morphology, and can be estimated easily.

(iii) The behavior of the spectra and the scattering outside the inertial range depends on turbule morphology, weakly for small K but strongly for large K.

(iv) In the inertial range, the spectra and the associated scattering cross sections are independent of turbule morphology. For example, it makes no difference whether the expectation turbule shape is Gaussian, exponential, or some other.

(v) If the spectra are chosen to match the Kolmogorov power law in the inertial range, then for fractal scaling, where ratios of successive sizes are the same for all sizes, turbule packing fractions must be size-independent, and temperature and velocity amplitudes must scale as the 1/3 power of the size.

(vi) Expressions for the structure constants (C_T^2,C_v^2) associated with the Kolmogorov spectra are obtained in terms of turbule model parameters and involve only scale-invariant expressions for fractal scaling.

(vii) For most ratios of outer scale length to wavelength and for some reasonable ratios of inner to outer scale length, the inertial range of scattering angles is significantly smaller than 0° to 180°, and, outside the inertial range, the scattering is significantly less than predicted by the full Kolmogorov spectrum.

(viii) For Gaussian model turbules, the structure functions may be expressed in terms of tabulated functions, and display the correct asymptotic behaviors as well as a 2/3-law dependence in an inertial range of distances determined by the inner and outer scales of turbulence.

Overall, then, we have accomplished the essential first step objective stated in the Introduction; to paraphrase it, we have ascertained that a turbule model of isotropic homogeneous turbulence works.

In order to predict the observed acoustic scattering by realistic models of turbulence, it is necessary to treat three things that were not considered in this work. One is a model of anisotropic inhomogeneous turbulence. In particular, anisotropy is probably different at different length scales, ranging from essentially isotropic at small scales to highly anisotropic at the largest scales. Inhomogeneity must be present because of the ground surface, but it should be maximum near the ground and decrease with altitude, and so again depends on turbulent length scales. Velocity turbules probably have correlated directions. A turbule model should be very helpful in representing these conditions.

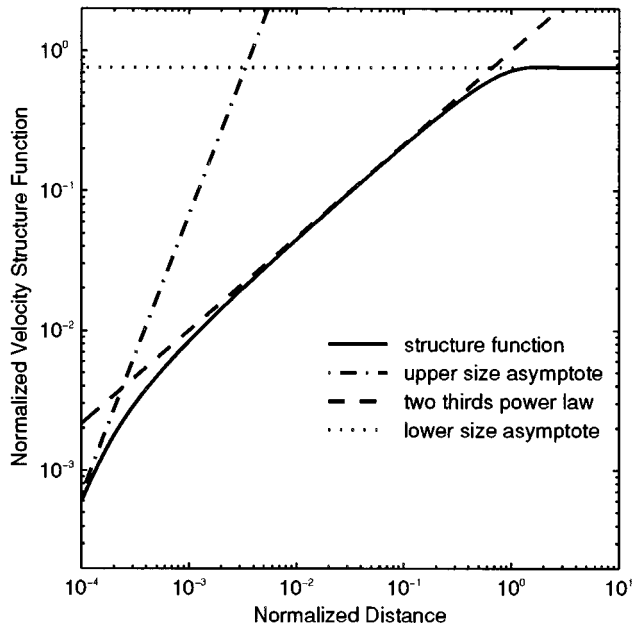


FIG. 18. Normalized structure function for an isotropic homogeneous Gaussian velocity turbule ensemble with $m=10^{-4}$ as a function of normalized distance r/a_1.

Another item is the effect of both the source and the detector actually being in the near field of much of the turbulence and having nonuniform radiation and acceptance patterns. Preliminary results indicate that a turbule model helps greatly in analyzing these effects, which are different for different length scales.

A third item is prediction of the magnitude of the fluctuations in the scattered intensity, compared to the mean scattered intensity. Experimentally, these fluctuations seem to be very large.[19] Preliminary calculations indicate that a turbule model is exceptionally well-suited to this analysis, and indeed predicts fluctuations with standard deviation at least equal to the mean.

APPENDIX. STRUCTURE FUNCTIONS

We show below that Eq. (77) reduces to Eq. (79) for the Gaussian envelope function (34). First, consider the integral in Eq. (77),

$$L_T(\xi) \equiv \int_0^\infty dz\ z^{-5/3}\left(1-\frac{\sin z}{z}\right)\int_{mz/\xi}^{z/\xi} dy\ y^{8/3}e^{-y^2/2}, \tag{A1}$$

where the second integral is $J_{8/3}^T(mz/\xi, z/\xi)$ for the Gaussian envelope function. Make a transformation to polar coordinates:

$$z=\rho\ \cos\ \theta, \quad y=\rho\ \sin\ \theta. \tag{A2}$$

Then from Eq. (A1), the limits on ρ are $(0,\infty)$, and the limits on θ are

$$\theta_1=\tan^{-1}(m/\xi), \quad \theta_2=\tan^{-1}(1/\xi). \tag{A3}$$

Then Eq. (A1) is

$$L_T(\xi)=\int_{\theta_1}^{\theta_2} d\theta\ \sin\ \theta(\tan\ \theta)^{5/3}[K_1(\theta)-K_2(\theta)], \tag{A4}$$

where

$$K_1(\theta)=\int_0^\infty d\rho\ \rho^2 e^{-(1/2)\rho^2\ \sin^2\ \theta}=\sqrt{\frac{\pi}{2}}\ \csc^3\ \theta, \tag{A5}$$

$$K_2(\theta)=\int_0^\infty d\rho\ \rho^2 e^{-(1/2)\rho^2\ \sin^2\ \theta}\left[\frac{\sin(\rho\ \cos\ \theta)}{\rho\ \cos\ \theta}\right]. \tag{A6}$$

In Eq. (A6), let $\eta=\cos\theta$. Then after an integration by parts,

$$K_2(\theta)=\tfrac{1}{2}(\cos\ \theta\ \sin^2\ \theta)^{-1}\int_0^\infty d\eta\ \cos\ \eta e^{-(1/2)\eta^2\ \tan^2\ \theta}. \tag{A7}$$

Writing $\cos\eta$ in terms of $e^{\pm i\eta}$ and completing the square in the exponent then yields

$$K_2(\theta)=\sqrt{\frac{\pi}{2}}\ \csc^3\ \theta\ \exp\left(-\frac{1}{2}\cot^2\ \theta\right). \tag{A8}$$

Using Eqs. (A5) and (A8) in Eq. (A4), letting $y=\cot\theta$, and using Eq. (A3) yields

$$L_T(\xi)=\sqrt{\frac{\pi}{2}}\int_\xi^{\xi/m} dy\ y^{-5/3}(1-e^{-y^2/2}). \tag{A9}$$

Using this, Eq. (78), and

$$J_{8/3}^T(0,\infty)=\int_0^\infty dy\ y^{8/3}e^{-y^2/2}=1.674 \tag{A10}$$

in Eq. (77) then yields Eq. (79).

The method for deriving Eq. (85) from Eq. (84) is identical with the above.

[1] V. I. Tatarskii, *The Effects of the Turbulent Atmosphere on Wave Propagation* (Keter, Jerusalem, 1971), Chap. 2.

[2] V. E. Ostashev, ''Sound propagation and scattering in media with random inhomogeneities of sound speed, density, and medium velocity,'' Waves Random Media **4**, 403–428 (1994).

[3] J. O. Hinze, *Turbulence* (McGraw-Hill, New York, 1975), 2nd ed., Chap. 3.

[4] V. E. Ostashev, ''Scattering of sound in randomly nonuniform moving medium and the possibility of sounding water-content fluctuation in the atmosphere,'' Atmos. Ocean. Phys. **27**, 956–961 (1991).

[5] A. S. Monin and A. M. Yaglom, *Statistical Fluid Mechanics: Mechanics of Turbulence* (MIT, Boston, 1975), Vol. 2, p. 385.

[6] R. J. Hill, ''Models of the scalar spectrum for turbulent advection,'' J. Fluid Mech. **88**, 541–562 (1978).

[7] V. Ostashev, V. Mellert, R. Wandelt, and F. Gerdes, ''Propagation of sound in a turbulent medium. I. Plane waves,'' J. Acoust. Soc. Am. (to be published).

[8] D. A. deWolf, ''A random motion model of fluctuations in a nearly transparent medium,'' Radio Sci. **18**, 138–142 (1983).

[9] W. E. McBride, H. E. Bass, R. Raspet, and K. E. Gilbert, ''Scattering of sound by atmospheric turbulence: Predictions in a refractive shadow zone,'' J. Acoust. Soc. Am. **91**, 1336–1340 (1992).

[10] A. S. Monin, ''Characteristics of the scattering of sound in a turbulent atmosphere,'' Sov. Phys. Acoust. **7**, 130 (1962).

[11] Most of these statements, and Eqs. (2), (3), (8), and (9) themselves, are not valid if the volume V_d observed by the detector does not include all of the illuminated part V_I of the volume V_T containing the turbulence. In such cases, violations of these equations may in fact be major, as we shall show in other work.

[12] The cross terms involving $\langle f_T f_v^* \rangle$ are zero because the temperature flux $\langle \Delta\tilde{T}_0(\mathbf{K})\tilde{v}_{0_2}^*(\mathbf{K})\rangle = 0$ for these conditions. However, this flux is nonzero and important for anisotropic turbulence.

[13] Our spectra Φ^T, Φ_{ij}^v are $(2\pi)^3(V_T)$ times the corresponding spectral densities ϕ, ϕ_{ij} of Ref. 1. Our C_v^2 is equal to $C^2\epsilon^{2/3}$ of that reference, where ϵ is the rate of energy dissipation per unit mass of fluid; our C_T^2 is equal to the C_T^2 of that reference.

[14] Note that this choice means that, for an isotropic ensemble of one-turbule systems, the average temperature variation of the turbule is not zero, so its forward temperature scattering is not zero either. In order to obtain zero average temperature variation for the turbulence overall, as required by Eq. (6), we treat collections of turbules constrained so that the sum of their average temperature variations is zero.

[15] S. Wolfram, *Mathematica* (Addison–Wesley, New York, 1991), 2nd ed.

[16] M. Nelkin, ''In what sense is turbulence an unsolved problem,'' Science **255**, 566–570 (1992).

[17] This is indeed (tacitly) required by the standard model, which gets $\nu = 1/3$ after stipulating that, in the cascade, the energy transfer rate per unit mass of the fluid is to be scale invariant.

[18] That is, typical inner scale lengths may be of the order of millimeters, while outer scale lengths may be of the order of tens to hundreds of meters.

[19] D. Havelock, M. R. Stinson, and G. A. Daigle, ''Phase and amplitude fluctuations in a refractive shadow,'' J. Acoust. Soc. Am. **92**, 2405(A) (1992).

Low Strouhal number instabilities of flow over apertures and wall cavities

M. S. Howe
Boston University, College of Engineering, 110 Cummington Street, Boston, Massachusetts 02215

(Received 4 March 1996; revised 17 April 1997; accepted 28 April 1997)

A theory is developed to predict the frequency of self-sustained oscillations of flow over an aperture in a plane wall and over a shallow wall cavity. The Mach number is sufficiently small that motion in the aperture or cavity may be regarded as incompressible. The shear layer thickness is small enough to permit it to be modeled by a vortex sheet, which is taken to be linearly disturbed from its equilibrium planar form. The motion of this sheet is discussed for circular and rectangular wall apertures, and numerical predictions are given for the Rayleigh conductivity $K_R(\omega)$ as a function of the radian frequency ω of the motion. Instabilities of the aperture flow are determined by poles of $K_R(\omega)$ in the upper complex frequency plane, and it is argued that the real parts of these complex frequencies correspond to the Strouhal numbers of self-sustained oscillations. An approximate method is given for determining the pole that corresponds to the minimum frequency, self-sustained oscillation. For incompressible flow there can be no net volume flux into a shallow wall cavity, and oscillations are in this case related to poles of a frequency-dependent drag coefficient. The predicted minimum Strouhal number for the cavity is close to measured values for the first stage of self-sustained oscillations of wall apertures and shallow cavities at very low Mach number. [S0001-4966(97)03608-4]

PACS numbers: 43.28.Py, 43.40.Rj, 43.50.Nm [LCS]

INTRODUCTION

Acoustic tones are frequently associated with instabilities of high-speed jets and shear layers. Unsteady pressures produced when a flow inhomogeneity perturbs the shear layer over the mouth of a wall cavity excite acoustic modes of the cavity, whose subsequent back reaction on the shear layer can lead to large amplitude oscillations which are sustained by the extraction of energy from the mean flow. An extensive discussion of these phenomena is given by Rockwell,[1] who also reviews research prior to 1983. Early notions of the excitation of cavity resonances attributed the oscillations to broadband excitation of acoustic modes by turbulence in the shear layer. However, self-sustained oscillations occur also when the flow is laminar, and it is often found that laminar flow resonances are the most intense.

Consider a wall cavity in the form of a rectangular cut out at depth l and length L in the streamwise direction. For shallow cavities ($L/l > 1$), flow excited tones do not generally correspond to acoustic modes of the cavity, and are not usually harmonically related. Strouhal numbers fL/U of tones of frequency f (U being the free stream velocity) typically lie within certain well-defined bands when plotted against mean flow Mach number.[2–5] The existence of these bands usually depends on *feedback*, according to which discrete vortices are formed periodically just downstream of the leading edge of the cavity, and proceed to convect towards the trailing edge. A sound pulse is generated by the interaction of a vortex with the trailing edge; the pulse propagates upstream, where its impingement on the leading edge induces separation of the boundary layer just upstream of the edge. A vortex travels across the cavity in time L/U_c, where the mean convection velocity $U_c \approx 0.4U-0.6U$, whereas the sound radiates back to the leading edge in time L/c_0. In the simplest approximation, therefore, the returning sound will arrive at the leading edge at just the right instant to reinforce the periodic shedding of vorticity provided the frequency f satisfies *Rossiter's equation*[2]

$$L/U_c + L/c_0 = n/f, \quad n = 1,2,3,\ldots, \tag{1}$$

where the index n is constant in each band or ''stage'' of the oscillation. When predictions of this equation are compared with experiment[5] it is found to be necessary to replace n by $n - \beta$, where β is a constant ''phase lag'' that depends on cavity depth; β/f is the aggregate time delay associated with (i) the arrival of a vortex at the trailing edge and the emission of the main acoustic pulse, and (ii) the arrival of the sound at the leading edge and the release of a new vortex.

Experiments at high subsonic Mach numbers[4,5] relate measured Strouhal numbers to the following modified Rossiter equation:

$$\frac{fL}{U} = (n-\beta) \Bigg/ \left(\frac{U}{U_c} + \frac{M}{\sqrt{[1+(\gamma-1)M^2/2]}} \right), \quad n = 1,2,3,\ldots, \tag{2}$$

where $M = U/c_0$ is the free stream Mach number based on the stagnation sound speed c_0 and γ is the ratio of specific heats. When $M > 0.2$, and for shallow, rectangular cavities with $L/l > 1$, predictions of Eq. (2) agree well with observations for $\beta \approx 0.25$ and $U_c/U \approx 0.6$, independently of the temperature of the free stream, the transverse width b of the cavity, and the Reynolds number UL/ν (ν being the kinematic viscosity).[5] These shallow cavity tones are governed by the feedback cycle discussed above; *cavity acoustic modes* (whose frequencies are determined by the interior cavity dimensions) tend to be unimportant unless $L/l < 2.5$

and M is large, when they can dominate the radiation provided the Strouhal number satisfies Eq. (2).

Theoretical models of shallow cavity tone excitation (e.g., Ref. 6) frequently postulate the existence of an acoustic ''source'' at the cavity trailing edge, whose strength and phase are estimated in terms of calculated properties of the shear layer motion near the edge. In flows of very low Mach number, however, when the wavelength of the sound is very much larger than the cavity mouth (when the mouth is ''acoustically compact''), the local flow can be regarded as incompressible, and the time delay L/c_0 in (1) from the returning sound becomes negligible.[7] For a deep cavity, the tones generally depend on feedback from standing acoustic waves within the cavity; the flow-acoustic coupling in such cases has been successfully modeled by a discrete vortex simulation of the unsteady flow over the mouth.[8–12]

This paper is concerned with the excitation of tonal sound by flow over an acoustically compact wall aperture or wall cavity when, in a first approximation, it is permissible to regard the local motion as incompressible. We do this by idealizing the shear layer as a vortex sheet that is linearly disturbed from its equilibrium form, and identifying tonal instabilities with disturbances of the sheet that grow exponentially with time. These disturbances are governed by the interaction of the vortex sheet with the solid surfaces and, for the wall aperture, are related to poles in the complex frequency plane of the Rayleigh conductivity $K_R(\omega)$ (the reciprocal of the acoustic impedance of the aperture).[13] An approximate method is used to determine the pole of $K_R(\omega)$ that is nearest to the real frequency axis. This corresponds to the lowest Strouhal number of tonal excitation [i.e., to $n=1$ in (1) and (2)]. The simplest case to treat by this means is one-sided grazing flow over an acoustically compact aperture in a thin, rigid wall. This is discussed in Sec. I for circular apertures (using calculated values of K_R for real ω given in Ref. 14) and in Sec. II for rectangular apertures. The method developed for rectangular apertures is extended in Sec. III to an aperture in a wall of finite thickness. In Sec. IV a similar procedure is applied to study self-sustained oscillations of flow over a shallow wall cavity.

I. SHEAR TONES GENERATED BY A CIRCULAR ORIFICE IN A THIN WALL

A. Conductivity of a circular aperture

Consider a small circular aperture in a thin rigid wall in the presence of a high Reynolds number, one-sided grazing flow at infinitesimal Mach number. Let the wall coincide with the plane $x_2=0$ of the rectangular coordinate system (x_1,x_2,x_3) with the center of the aperture at the origin, and suppose the flow is confined to the ''upper'' region $x_2>0$ (see Fig. 1) at speed U in the x_1 direction. The shear layer over the aperture will be assumed to be linearly perturbed by a uniform, time-dependent pressure $p_0(t)$ applied in $x_2>0$. This produces a volume flux $Q(t)$ through the aperture (taken as positive in the $+x_2$ direction), as a result of which sound waves are radiated into the fluid on either side of the wall. The influence of mean flow on sound propagation may be ignored provided the Mach number $M\equiv U/c_0$ is small

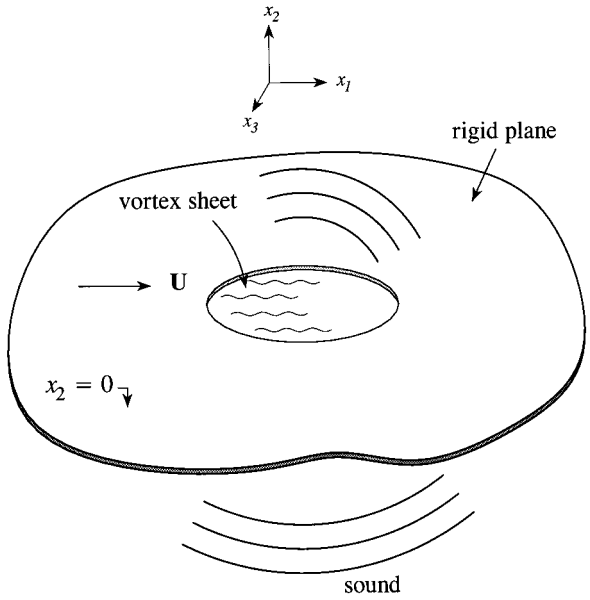


FIG. 1. Flow over a circular aperture in a thin, rigid wall.

(c_0 being the speed of sound).[15] Attention is here confined to low frequency motion, such that the acoustic wavelength is much larger than the aperture diameter, so that the acoustic pressure at large distances $|\mathbf{x}|$ from the aperture is the *monopole* field

$$p_Q=\pm\frac{\rho_0}{2\pi|\mathbf{x}|}\frac{\partial Q}{\partial t}\left(t-\frac{|\mathbf{x}|}{c_0}\right), \quad \text{for} \quad x_2\gtrless 0. \tag{3}$$

The flux Q can be expressed in terms of the applied pressure p_0 by introducing the Rayleigh conductivity K_R of the aperture.[13] This is usually defined for time-harmonically varying quantities (proportional to $e^{-i\omega t}$) by the relation

$$K_R(\omega)=i\omega\rho_0 Q(\omega)/[p(\omega)], \tag{4}$$

where ρ_0 is the mean fluid density, $Q(\omega)\equiv(1/2\pi)\int_{-\infty}^{\infty}Q(t)e^{i\omega t}\,dt$ is the Fourier component of the flux Q of radian frequency ω, and $[p]\equiv p_+-p_-$ is the difference in the time harmonic pressures above and below the wall that forces the fluid through the aperture. In the present case $p_-=0$ and $p_+(\omega)=p_0(\omega)$.

In the time domain Eq. (4) must be expressed in terms of a convolution integral:

$$\rho_0\frac{\partial Q(t)}{\partial t}=-\int_{-\infty}^{\infty}K_R(\omega)p_0(\omega)e^{-i\omega t}\,d\omega. \tag{5}$$

It may be assumed that $p_0(t)$ vanishes prior to some initial time t_0, say, and therefore that $p_0(\omega)\to 0$ as Im $\omega\to+\infty$ and is regular in the upper half of the complex frequency plane; $p_0(\omega)$ will be regular *everywhere* if the applied pressure $p_0(t)$ is of finite duration.[16] To ensure that $Q(t)=0$ for $t<t_0$ the path of integration in (5) must pass above any singularities of $K_R(\omega)$. For $t>t_0$ the integral can be evaluated by displacing the integration contour downwards towards the real axis, ''wrapping'' it around any singularities of the integrand in Im $\omega>0$. The contributions from these

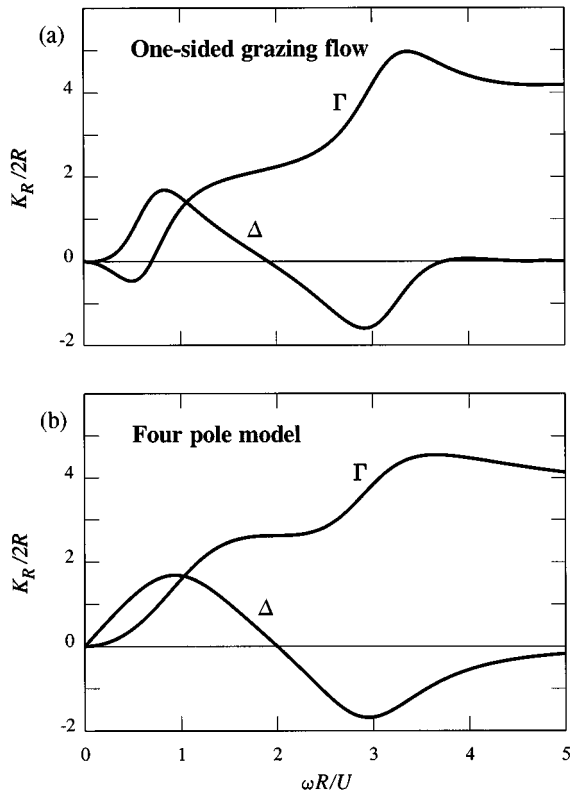


FIG. 2. (a) Normalized conductivity $K_R/2R=\Gamma-i\Delta$ of a circular aperture in a thin wall calculated by Scott (Refs. 14 and 17). (b) Four-pole approximation.

singularities (if they exist) will dominate the integral as $t-t_0\to+\infty$.

In the absence of mean flow the conductivity $K_R=2R$, where R is the aperture radius.[13] The conductivity in the presence of a one-sided grazing flow, when the shear layer is modeled by a vortex sheet as in Fig. 1, has been computed by Scott[14,17] for real frequencies ω. Scott assumed the vortex sheet to be linearly disturbed by the pressure differential $[p]$, and solved numerically the integrodifferential equation satisfied by its displacement $\zeta(x_1,x_3)e^{-i\omega t}$ (in the x_2 direction), subject to the Kutta condition that the sheet leaves the upstream semicircular edge of the aperture tangentially.[18] K_R is now complex, and varies significantly with the Strouhal number $\omega R/U$. The real and imaginary parts of $K_R(\omega)/2R\equiv(\Gamma-i\Delta)$ are plotted in Fig. 2(a) against $\omega R/U$ for real values of ω. When $\Delta\neq0$ and ω is real, the interaction of the applied pressure $p_0(\omega)$ with the aperture is nonconservative, in the sense that there is an irreversible transfer of perturbation energy to or from the mean flow according as $\Delta\gtrless0$.

B. Analytic continuation of $K_R(\omega)$

Since $Q(t)$ and $[p(t)]$ are real valued quantities, the conductivity satisfies $K_R(-\omega^*)=K_R^*(\omega)$, where the asterisk denotes complex conjugate. For *real* values of ω we have

$$\Gamma(-\omega)=\Gamma(\omega),\quad \Delta(-\omega)=-\Delta(\omega). \tag{6}$$

If the motion were stable, $K_R(\omega)/2R\equiv\Gamma(\omega)-i\Delta(\omega)$would be regular in Im $\omega>0$. In general $\Gamma(\omega)\to\Gamma_\infty=$const as $|\omega|\to\infty$ and $\Delta(\omega)\to0$ [cf. Fig. 2(a)]. Then (for stable motion) $f(\omega)\equiv\Gamma(\omega)-\Gamma_\infty-i\Delta(\omega)$ is regular in the upper half-plane and vanishes as $|\omega|\to\infty$. Cauchy's theorem[16] applied to a closed contour consisting of the real axis and a semicircle in the upper half-plane whose radius grows without limit, implies that $f(\omega)=(1/2\pi i)\int_{-\infty}^{\infty}f(\xi)d\xi/(\xi-\omega)$ for Im$(\omega)>0$, where the integration is along the real axis. When ω approaches the real axis, the real and imaginary parts of this equation and (6) yield the *Kramers–Kronig* relations[19]

$$\begin{aligned}\Gamma(\omega)-\Gamma_\infty&=-\frac{2}{\pi}\int_0^\infty\frac{\xi\Delta(\xi)d\xi}{\xi^2-\omega^2},\\ \Delta(\omega)&=\frac{2\omega}{\pi}\int_0^\infty\frac{[\Gamma(\xi)-\Gamma_\infty]d\xi}{\xi^2-\omega^2}\quad(\omega\text{ real}),\end{aligned} \tag{7}$$

where the integrals are principal values.

Alternatively, when $\Gamma(\omega)$ and $\Delta(\omega)$ have been determined by experiment or numerical computation, Eqs. (7) can be used to investigate the stability of the motion. If the equations are not satisfied $K_R(\omega)$ must be singular in Im $\omega>0$, and the motion is *absolutely unstable*, in that the smallest perturbation of the flow can cause a spontaneous growth of large amplitude motions in the aperture. This happens to be the case for the circular aperture with one-sided flow studied by Scott: the Kramers–Kronig formulas (7) are not satisfied by the numerically derived values of Γ and Δ shown in Fig. 2(a). More information concerning the nature of the unstable motion can be derived by expressing $f(\omega)=\Gamma(\omega)-\Gamma_\infty-i\Delta(\omega)$ for real values of ω as the sum $f^+(\omega)+f^-(\omega)$ of two functions which are, respectively, regular in Im $(\omega)\gtrless0$. We can do this by making use of the following extension of (7):[20]

$$\begin{aligned}\Gamma^\pm(\omega)-\Gamma_\infty^\pm&=\frac{\mp1}{2\pi}\int_{-\infty}^{\infty}\frac{\Delta(\xi)d\xi}{\xi-\omega}+\frac{1}{2}[\Gamma(\omega)-\Gamma_\infty],\\ \Delta^\pm(\omega)&=\frac{\pm1}{2\pi}\int_{-\infty}^{\infty}\frac{[\Gamma(\xi)-\Gamma_\infty]d\xi}{\xi-\omega}+\frac{1}{2}\Delta(\omega),\end{aligned} \tag{8}$$

where ω is real.

These formulas can be used to derive an approximate analytic continuation of $K_R(\omega)$ into the complex frequency plane. In several important cases it is permissible to assume that all singularities of the conductivity are simple poles occurring in pairs at

$$\kappa\equiv\omega R/U=\pm\kappa_n-i\beta_n\quad(n=1,2,\ldots),$$

where κ_n and β_n are real. We then write

$$\begin{aligned}\Gamma(\omega)-i\Delta(\omega)=&\sum_{n=1}^{N}\left(\frac{\alpha_n}{\kappa-(\kappa_n-i\beta_n)}-\frac{\alpha_n^*}{\kappa+(\kappa_n-i\beta_n)}\right)\\&+\sum_{n=1}^{N}\frac{2(\alpha_{rn}\kappa_n-\alpha_{in}\beta_n)}{\kappa_n^2+\beta_n^2},\end{aligned} \tag{9}$$

where α_n are complex constants with real and imaginary parts, respectively, equal to α_{rn} and α_{in}. When $\Gamma^\pm(\omega)$ and $\Delta^\pm(\omega)$ have been calculated for real ω from Eqs. (8), the coefficients in (9) are chosen to give the best fit to $\Gamma^\pm(\omega)-i\Delta^\pm(\omega)$ using those terms in the summation correspond-

ing, respectively, to poles in the lower and upper halves of the ω plane.

This approximation has been used to extend the numerical results of Scott[17] into the complex frequency domain, by means of the four-pole model defined by

$$\alpha_1=1.23,\ \kappa_1=0.9,\ \beta_1=1;\ \alpha_2=0.82,\ \kappa_2=2.9,$$
$$\beta_2=-0.8;\ \alpha_n=0,\ n>2. \tag{10}$$

There are two poles in the upper half-plane at $\omega R/U=\pm 2.9+0.8i$ and two in the lower half-plane at $\omega R/U=\pm 0.9-i$. The real and imaginary parts of $K_R(\omega)/2R$ defined by this approximation are plotted in Fig. 2(b) for comparison with Scott's numerical results in Fig. 2(a).

C. Self-sustained aperture tones

When the approximation for $K_R(\omega)$ defined by Eqs. (9) and (10) is used to evaluate the integral (5), the aperture flux at large times is determined by residue contributions from the poles in the upper half-plane, which yield a monopole source strength that increases exponentially with time, i.e.,

$$\rho_0\frac{\partial Q}{\partial t}\approx -8\pi U\alpha_2|p_0(\Omega)|\cos\left\{\frac{2.9Ut}{R}-\varphi\right\}e^{0.8Ut/R},$$

where $\Omega R/U\equiv 2.9+0.8i$ and $\arg\{p_0(\Omega)\}=\varphi$ [note that $p_0(-\Omega^*)\equiv p_0^*(\Omega)$ when $p_0(t)$ is real]. In practice, of course, this exponential growth is curtailed by nonlinear mechanisms ignored by linear perturbation theory. However, since according to (1) the frequency of the motion is controlled by the convection velocity across the aperture, which does not depend significantly on amplitude, the frequency of nonlinear oscillations would be expected to be close to the linear theory value. Because of the exponential growth, this tonal component of the sound can in practice be initiated by any small disturbance and appears to be spontaneously excited by the flow. When the frequency is expressed as a *Strouhal* number fD/U, where $f=\omega/2\pi$ and $D=2R$ is the aperture diameter, the above calculation suggests that the lowest tone occurs near $fD/U\approx 0.9$. When the Strouhal number is based on the average length of the aperture in the streamwise direction, $L=(\pi/4)D$, we find $fL/U\approx 0.7$. This value is close to Strouhal numbers observed experimentally, although most measurements have been made using cavity backed apertures beneath relatively thick shear layers, where the shear layer motion is coupled to an acoustic mode of the cavity and for which $fL/U\approx 0.3-0.5$, but larger values have been reported at very small mean flow speeds.[4,7,10,12,21,22]

II. RECTANGULAR ORIFICE IN A THIN WALL

A. The Rayleigh conductivity

The numerical method of Scott[14,17] is readily extended to determine the influence of grazing flow (in the vortex sheet approximation) on the conductivity of an orifice of arbitrary shape, and in particular for the important case of a rectangular aperture. In this section, however, we shall outline a simpler, approximate numerical scheme (discussed in detail in Ref. 23) for a rectangular aperture that is easily generalized to cases where the wall is of finite thickness, or to finite depth cut outs with rectangular openings.

The configuration is identical to that of Sec. I, except that the aperture is rectangular with sides of lengths L and b, respectively, parallel and transverse to the mean flow. Let the aperture occupy the region $|x_1|<s\equiv L/2$, $|x_3|<\frac{1}{2}b$. When a time harmonic pressure differential $[p]=p_+-p_-$ (proportional to $e^{-i\omega t}$) is applied across the wall, the resulting (incompressible) unsteady motion above and below the aperture can be expressed in terms of velocity potentials

$$\varphi^{\pm}(\mathbf{x})=\frac{\mp 1}{2\pi}\int\frac{v_2(y_1,\pm 0,y_3)}{|\mathbf{x}-\mathbf{y}|}dy_1\,dy_3$$
$$(y_2=0,\ x_2\gtrless 0), \tag{11}$$

where the integration is over the plane of the wall, and v_2 is the x_2 component of velocity (normal to the wall). This velocity vanishes at the wall except in the aperture where, in the linearized approximation,

$$v_2(x_1,x_2,x_3)\to\left(-i\omega+U\frac{\partial}{\partial x_1}\right)\zeta,\quad -i\omega\zeta,$$
$$\text{according as } x_2\to\pm 0, \tag{12}$$

$\zeta(x_1,x_3)$ being the displacement of fluid particles (in the x_2 direction) which lie in the plane of the wall in the undisturbed state.

The equation of motion of the vortex sheet is obtained by equating the pressures $p_\pm+p'_\pm$ on opposite sides of the vortex sheet, where $p'_\pm$ denotes the pressure fluctuations above and below the wall produced by motion of the sheet. These pressures are calculated from (11), (12), and the linearized form of Bernoulli's equation, whereupon we find that ζ satisfies

$$\left[\left(\omega+iU\frac{\partial}{\partial x_1}\right)^2+\omega^2\right]\int_s\frac{\zeta(y_1,y_3)}{2\pi|\mathbf{x}-\mathbf{y}|}dy_1\,dy_3$$
$$=[p]/\rho_0,\quad x_2=y_2=0, \tag{13}$$

where $\mathbf{x}\equiv(x_1,0,x_3)$ and the integration is over the surface S of the aperture.

We assume that vortex shedding from the upstream edge $(x_1=-s)$ of the aperture produces strongly correlated motion within S at different transverse locations x_3, such that ζ may be regarded as independent of x_3. The y_3 integration in (13) may then be performed explicitly. If we also integrate with respect to x_3, and introduce the notation

$$\frac{-1}{2b}\int\int_{-b/2}^{b/2}\frac{dy_3\,dx_3}{\sqrt{(x_1-y_1)^2+(x_3-y_3)^2}}$$
$$=\ln|\xi-\eta|+\mathscr{S}(\xi,\eta),\quad \xi=x_1/s,\quad \eta=y_1/s, \tag{14}$$

where

$$\mathscr{S}(\xi,\eta)=-\ln\{b/s+\sqrt{[(b/s)^2+(\xi-\eta)^2]}\}$$
$$+\sqrt{[1+(s/b)^2(\xi-\eta)^2]}-(s/b)|\xi-\eta|, \tag{15}$$

then (13) becomes

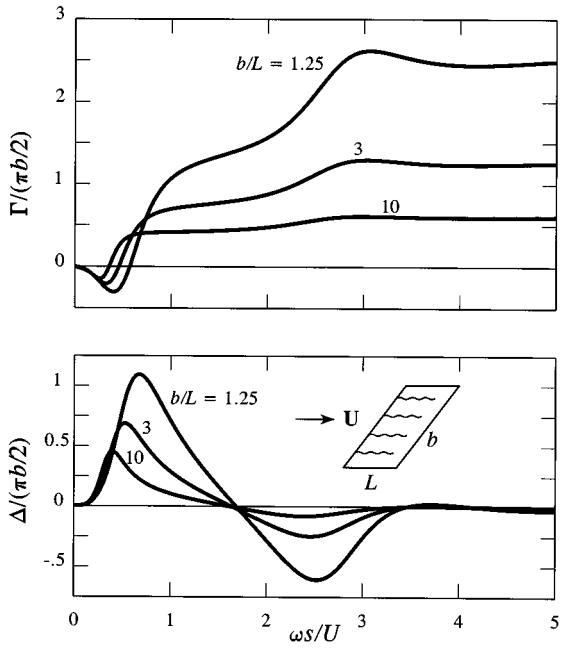


FIG. 3. Conductivity $K_R=\pi/2b(\Gamma-i\Delta)$ of a rectangular aperture of different aspect ratios b/L in a thin wall.

$$\left[\epsilon^2+\left(\epsilon+i\frac{\partial}{\partial\xi}\right)^2\right]\int_{-1}^{1}\zeta(\eta)\{\ln|\xi-\eta)|+\mathscr{L}(\xi,\eta)\}d\eta$$

$$=-\pi s[p]/\rho_0U^2,\quad |\xi|<1, \tag{16}$$

where $\epsilon=\omega s/U$.

This equation is integrated with respect to the second-order differential operator on the left-hand side. By setting

$$\zeta'=-(2\rho_0\omega^2 s/\pi[p])\zeta,$$

we find

$$\int_{-1}^{1}\zeta'(\eta)\{\ln|\xi-\eta|+\mathscr{L}(\xi,\eta)\}d\eta+\lambda_1e^{i\sigma_1\xi}+\lambda_2e^{i\sigma_2\xi}$$

$$=1,\quad |\xi|<1, \tag{17}$$

where $\sigma_1=(\omega s/U)(1+i)$, $\sigma_2=(\omega s/U)(1-i)$ are the nondimensional Kelvin–Helmholtz wave numbers for a vortex sheet,[24] and λ_1 and λ_2 are constants of integration.

The integral equation can be solved by collocation. The values of λ_1 and λ_2 are chosen to ensure that the Kutta condition is satisfied at the upstream edge of the aperture, i.e., that $\zeta=\partial\zeta/\partial\xi=0$ at $\xi=-1$. The numerical solution and the definition (4) then supply the Rayleigh conductivity from the formula

$$K_R=-\frac{\pi}{2}b\int_{-1}^{1}\zeta'(\eta)d\eta. \tag{18}$$

Typical predictions of $K_R\equiv(\pi/2)b(\Gamma-i\Delta)$ are plotted in Fig. 3 for the three aspect ratios $b/L=1.25$, 3, and 10. The results for $b/L=10$ are in close agreement with an asymptotic formula for large b/L given in Ref. 14.

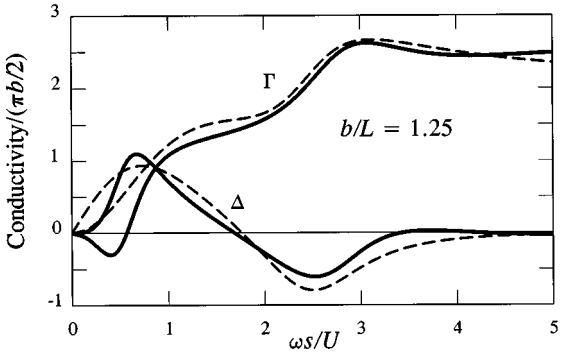


FIG. 4. Comparison of the four-pole approximation (19) (----) with the calculated conductivity (———) for $b/L=1.25$.

B. Aperture tones

There is a strong qualitative similarity between the predictions shown in Figs. 2 and 3 for circular and rectangular apertures. Using the method of Sec. I B we can develop approximate analytic continuations of K_R into the complex frequency domain. In Fig. 4 a four pole approximation of the type (9) (in which κ is replaced by $\omega s/U$) is compared to the numerically predicted conductivity for the case $b/L=1.25$, where

$$\alpha_1=2.4,\ \kappa_1=0.5,\ \beta_1=1.15;\quad \alpha_2=0.75,\ \kappa_2=2.45,$$
$$\beta_2=-0.7;\quad \alpha_n=0,\quad n>2. \tag{19}$$

According to the argument of Sec. I C, the poles at $\omega s/U=\pm2.45+0.7i$, correspond to an instability of the aperture motion that is manifested by shear layer oscillations at a frequency close to $\omega s/U=2.45$. In a first approximation, this frequency coincides with that at which Δ assumes its minimum value in Fig. 4. Figure 3 indicates that this minimum moves to marginally lower frequencies as the aspect ratio b/L increases, although the variation is small, and the minimum occurs near $\omega s/U=2.3$ when b/L becomes large. It may also be noted that, when the characteristic Strouhal number $\omega R/U\approx2.9$ obtained in Sec. I for the circular aperture is scaled by replacing the length R by the mean semichord $\overline{s}=(\pi/4)R$ of the circular aperture, we find $\omega\overline{s}/U\approx2.28$, which is close to the rectangular aperture case when b/L exceeds about 3.

III. RECTANGULAR APERTURE IN A THICK WALL

A. Approximate representation of the conductivity

The method of Sec. II for a rectangular aperture is now extended to include the effect of wall thickness on the conductivity. We discuss here the simplest case in which the wall thickness h (see Fig. 5) is large enough for the coupling between the motion in the upper mouth (spanned by the vortex sheet) and the lower mouth (where there is no mean flow) to be expressed in terms of the locally uniform, time harmonic pressures p_U and p_L produced by the motion, respectively, just below the vortex sheet and just above the lower open mouth. Referring to Fig. 5, if Q is the volume flux

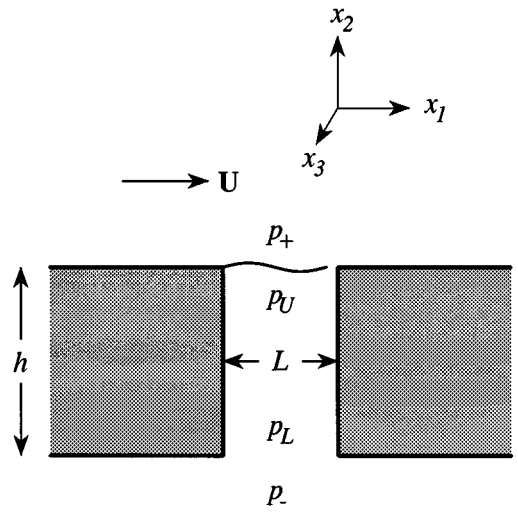


FIG. 5. One-sided grazing flow over a rectangular aperture in a thick wall.

through the aperture (in the $+x_2$ direction), the conductivities K_U, K_L of the upper and lower mouths, respectively, satisfy

$$p_+ - p_U = i\rho_0\omega Q/K_U, \quad p_- - p_L = -i\rho_0\omega Q/K_L, \tag{20}$$

where the minus sign on the right of the second of these equations occurs because K_L is defined with respect to a flux $-Q$ that is directed out of the mouth. In the same approximation, the motion of the cylindrical column of fluid within the aperture satisfies the momentum equation

$$A(p_L - p_U) = -i\omega h\rho_0 Q, \tag{21}$$

where $A = bL$ is the aperture area. Combining these equations, the net conductivity K_R of the aperture in the thick wall is given by

$$\frac{1}{K_R} = \frac{1}{K_U} + \frac{1}{K_L} + \frac{h}{A}. \tag{22}$$

The length A/K_L is the ''end-correction'' of the lower mouth.[13] The presence of the vortex sheet causes the end-correction A/K_U of the upper mouth to be complex. The reciprocal $F_R = 1/K_R$ is the aperture acoustic impedance, in terms of which (22) assumes the simpler form

$$F_R = F_U + F_L + h/A, \tag{23}$$

where F_U and F_L are the acoustic impedances of the upper and lower mouths, respectively.

B. Determination of F_U

Take the coordinate origin in the plane of the upper surface of the wall at the center of the mouth. Assume, as in Sec. II, that the spanwise variation of the displacement ζ of the vortex sheet can be ignored. When the pressure just above the vortex sheet is averaged over the span, we find

$$p = p_+ + \frac{\rho_0 U^2}{\pi s}\left(\epsilon + i\frac{\partial}{\partial\xi}\right)^2 \int_{-1}^{1} \zeta(\eta)\{\ln|\xi-\eta| + \mathscr{L}(\xi,\eta)\}d\eta, \quad |\xi| \equiv |x_1/s| < 1, \tag{24}$$

where $\mathscr{L}(\xi,\eta)$ is defined as in (15).

The motion in the cylindrical region below the vortex sheet is uniform in the spanwise (x_3) direction. If the vortex sheet is temporarily replaced by a rigid lid, the local incompressible flow produced by a spanwise line source of unit strength at $x_1 = y_1$, $x_2 = -0$, $|x_3| < b/2$ can be calculated by making use of the two-dimensional velocity potential for a source injecting fluid into a semi-infinite, uniform duct, which can be calculated by the method of conformal transformation.[24] If the mean pressure just below the vortex sheet is p_U, the pressure on the lower surface of the vortex sheet can then be written

$$p = p_U - \frac{\rho_0 U^2 \epsilon^2}{\pi s}\int_{-1}^{1} \zeta(\eta)\{\ln|\xi-\eta| + \mathscr{L}_U(\xi,\eta)\}d\eta, \quad |\xi| < 1, \tag{25}$$

where

$$\mathscr{L}_U(\xi,\eta) = \ln\left(\frac{4\sin\{\pi(\xi-\eta)/4\}\cos\{\pi(\xi+\eta)/4\}}{\xi-\eta}\right).$$

Equating the pressures (24) and (25) and rearranging, the equation of motion of the vortex sheet becomes

$$\left[\epsilon^2 + \left(\epsilon + i\frac{\partial}{\partial\xi}\right)^2\right]\int_{-1}^{1}\zeta(\eta)\{\ln|\xi-\eta| + \mathscr{L}(\xi,\eta)\}d\eta + \epsilon^2\int_{-1}^{1}\zeta(\eta)\{\mathscr{L}_U(\xi,\eta) - \mathscr{L}(\xi,\eta)\}d\eta = -\pi s(p_+ - p_U)/\rho_0 U^2, \quad |\xi| < 1. \tag{26}$$

Proceeding as in the treatment of Eq. (16), we find

$$\int_{-1}^{1}\zeta'(\eta)\{\ln|\xi-\eta| + \mathscr{L}(\xi,\eta) + \mathfrak{C}(\xi,\eta)\}d\eta + \lambda_1 e^{i\sigma_1\xi} + \lambda_2 e^{i\sigma_2\xi} = 1, \quad |\xi| < 1, \tag{27}$$

where $\zeta' = -\{2\rho_0\omega^2 s/\pi(p_+ - p_U)\}\zeta$, and

$$\mathfrak{C}(\xi,\eta) = \frac{\epsilon}{2}\int_{-1}^{1}\{\mathscr{L}_U(\mu,\eta) - \mathscr{L}(\mu,\eta)\} \times \exp\{i\epsilon(\xi-\mu) - \epsilon|\xi-\mu|\}d\mu.$$

Equation (27) is now solved by collocation, as before (with the Kutta condition imposed at $\xi = -1$), and the solution used to determine $F_U \equiv 1/K_U$ from the relation (18). Typical predictions for the aspect ratios $b/L = 1.25$, 3, and 10 considered previously, are shown in Fig. 6.

C. Determination of F_L

The acoustic impedance F_L of the lower mouth can be found by the same procedure. There is now no vortex sheet, and the corresponding normalized displacement ζ' of fluid in the plane of the mouth in the undisturbed state is obtained by

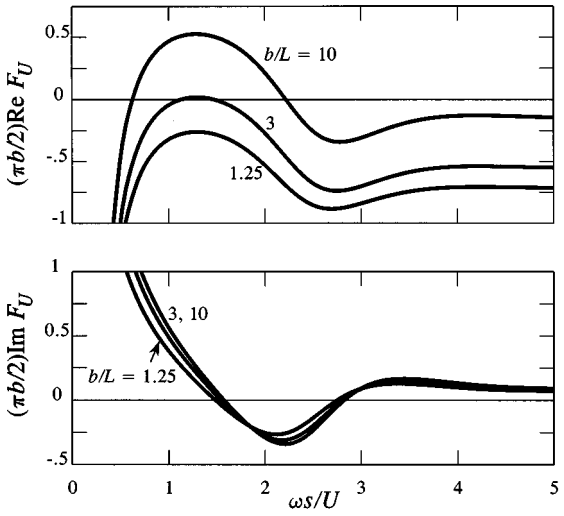


FIG. 6. Vortex sheet approximation to the acoustic impedance F_U of a rectangular opening in a thick wall in the presence of a grazing flow.

equating expressions for the pressure at corresponding points just above and just below the level $x_2=-l$ of the mouth. Because there is no mean flow this yields a governing equation that does not involve differential operators:

$$\int_{-1}^{1} \zeta'(\eta)(\ln|\xi-\eta|+\tfrac{1}{2}\{\mathscr{S}(\xi,\eta)+\mathscr{S}_U(\xi,\eta)\})\,d\eta=1, \tag{28}$$

where $F_L \equiv 1/K_L$ is determined by (18). The Kutta condition cannot be imposed in this case, and ζ' exhibits mild singularities at the ends $\xi=\pm 1$ of the mouth.

The variations of F_L and K_L with aspect ratio b/L are shown in Fig. 7. Note that, the present locally two-dimensional approximation which yields the one-dimensional integral equation (28) might have been expected to be applicable only for large aspect ratios, since in the absence of flow there is no reason to prefer an equation averaged over the spanwise as opposed to the chordwise dimension. However, inspection of Fig. 7 reveals that for $0.5 < b/L < 2.5$, K_L is well approximated by Rayleigh's formula[13]

$$K_L \approx 3.83\sqrt{A/\pi}, \quad A=bL \tag{29}$$

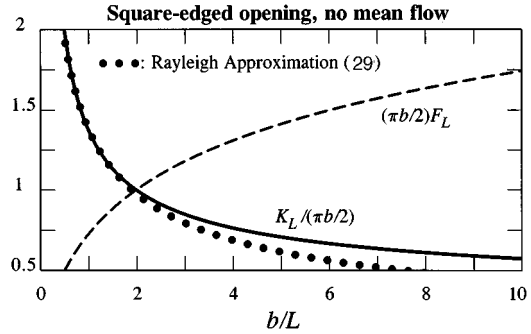


FIG. 7. Predicted dependence on aspect ratio b/L of the conductivity K_L and impedance F_L of a rectangular opening in a wall in the absence of mean flow.

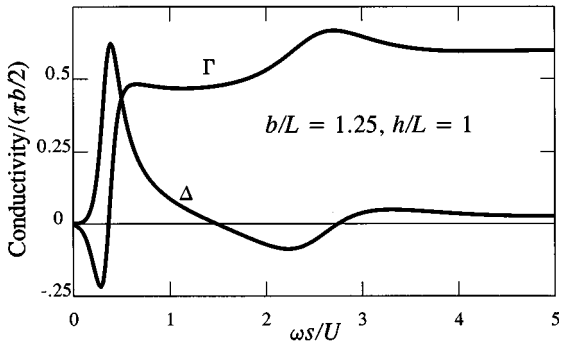


FIG. 8. Predicted conductivity $K_R=(\pi/2)b(\Gamma-i\Delta)$ of a rectangular aperture in a thick wall with one-sided grazing flow for $b/L=1.25$, $h/L=1$.

(shown dotted in the figure) which is applicable for any, nonelongated opening of area A in a wall. At larger aspect ratios $K_L \approx 2\pi b/\ln\{8b/Le\}$, where $e\approx 2.178$ is the exponential constant.[14]

D. Flow excited aperture tones

These numerical results may now be used in (22) to calculate the conductivity $K_R \equiv (\pi/2)b(\Gamma-i\Delta)$ of the wall aperture for different wall thickness ratios h/L. Figure 8 illustrates the result of such a calculation for $h/L=1$ and $b/L=1.25$. Both curves in this figure are qualitatively similar to those shown in Fig. 4 for a thin wall aperture of the same aspect ratio. The overall reduction in the magnitudes of Γ and Δ relative to those in Fig. 4 is a consequence of the thickness contribution h/A to the acoustic impedance. In a first approximation, the argument of Sec. I B implies that the frequency at which Δ assumes its minimum (and negative) value may be taken to correspond to the minimum frequency of self-sustained oscillations of the flow over the aperture. On this basis we can calculate the dependence of this frequency on wall thickness for different aspect ratios, as shown in Fig. 9 for the three cases $b/L=1.25$, 3, and 10, considered previously.

The results are expressed as a Strouhal number fL/U $(f=\omega/2\pi)$. According to Fig. 9, fL/U decreases with increasing wall thickness, becoming ultimately constant and

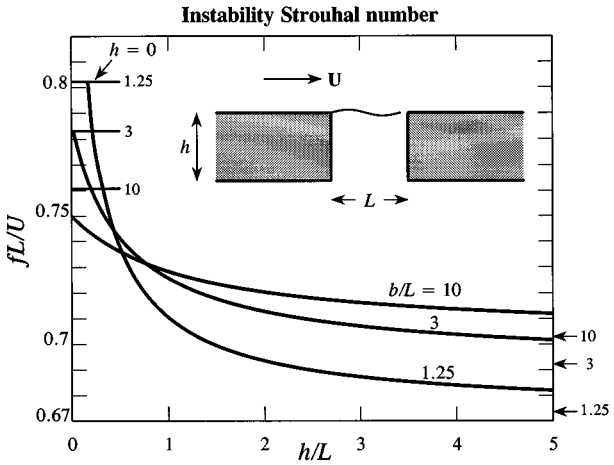


FIG. 9. Calculated Strouhal number of self-sustaining oscillations for one-sided grazing flow over a rectangular aperture in a wall of thickness h.

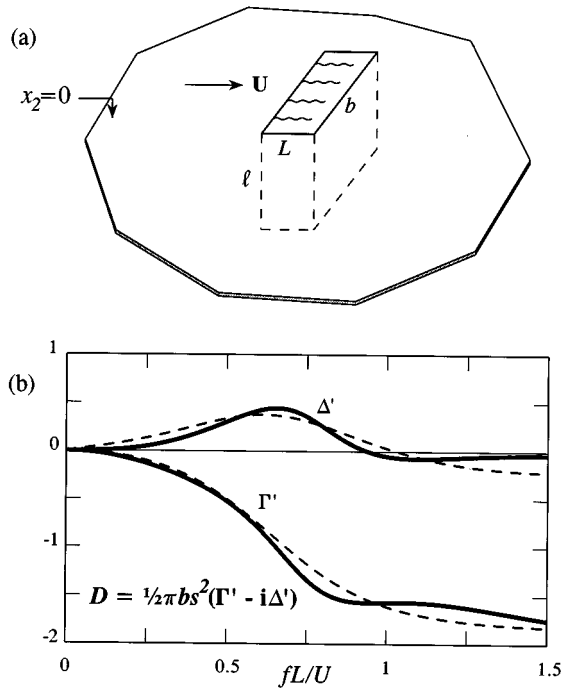


FIG. 10. (a) Rectangular wall cavity. (b) Real and imaginary parts of the function $D(\omega)$ of Eq. (31) for a wall cavity with $b/L=1$, and the four-pole approximation (----).

equal to asymptotic values indicated on the right of the figure. Recall, however, that the present calculation assumes the motion within and in the immediate neighborhood of the aperture to be effectively incompressible, so that the wall thickness h must always be small compared to the characteristic wavelength $\approx L/M$ $(M=U/c_0)$of sound of frequency f. The present calculation is not applicable when h is small compared to L; also shown in the figure are the limiting values as $h/L\to 0$ (labeled ''$h=0$'')predicted by the thin wall theory of Sec. II. The numerical results for $h/L\neq 0$ are seen to increase steadily towards these respective limiting values as h/L approaches zero. The characteristic Strouhal numbers are typically of order 0.7 which is consistent with Strouhal numbers observed for shallow cavities at very low Mach numbers.[5]

IV. THE SHALLOW WALL CAVITY

At low Mach numbers the problem of sound generation by flow over a wall cavity whose depth l is very much smaller than the acoustic wavelength is similar to that considered above for an orifice in a rigid wall. Most experimental investigations have been conducted at $M=U/c_0>0.2$, and are relevant to vibration and buffeting problems encountered by aircraft.[15]

At low frequencies the cavity might be expected to behave like a Helmholtz resonator, with a monopole radiation field of the type (3). At very low frequencies, however, the fluid in the cavity is effectively *incompressible*, the volume flux $Q\to 0$, and the monopole strength ultimately vanishes. To fix ideas, consider the excitation of low-frequency oscillations by low Mach number flow over a square edged cut out illustrated in Fig. 10(a), and take the coordinate origin in the center of the mouth. In the limit $M\equiv U/c_0\to 0$ the acoustic pressure at large distances from the cavity can be expressed in the form

$$p(\mathbf{x},t)\approx \frac{\rho_0}{2\pi|\mathbf{x}|}\frac{\partial}{\partial t}\int_S v_2(\mathbf{y},t-|\mathbf{x}-\mathbf{y}|/c_0)\,dy_1\,dy_3 \quad \text{as } |\mathbf{x}|\to\infty, \tag{30}$$

where the integration is over the mouth S where the normal velocity is v_2. For an acoustically compact opening, when the flux $Q\equiv 0$, the first nontrivial term in the expansion of the integrand in powers of y/c_0 gives the dipole field

$$p(\mathbf{x},t)\approx \frac{x_j}{2\pi c_0|\mathbf{x}|^2}\frac{\partial F_j}{\partial t}(t-|\mathbf{x}|/c_0),\quad |\mathbf{x}|\to\infty,$$

where $F_j=\rho_0\int_S y_j\partial v_2(\mathbf{y},t)/\partial t\,dy_1\,dy_3$ $(j=1$ or 3) is the unsteady *drag* exerted on the fluid by the cavity.

Now at very low frequencies, a uniform, time-dependent pressure p_0 applied to the cut out in $x_2>0$ cannot produce drag. We must therefore consider an applied tangential pressure force $-\partial p_0/\partial x_j$. Suppose the pressure gradient is in the mean flow direction (the x_1 direction), and that its spatial variation can be neglected. Then by analogy with (5) we can write

$$F_1(t)=-\int_{-\infty}^{\infty} D(\omega)\partial p_0(\omega)/\partial x_1 e^{-i\omega t}\,d\omega, \tag{31}$$

where the *drag coefficient* $D\equiv M_{11}/\rho_0$, M_{11} being the effective ''added mass'' of the cavity in the presence of the shear layer.[24] The frequencies of self-sustained radiation from the cut out are determined by the singularities of $D(\omega)$ in the upper half of the ω plane.

The drag coefficient can be calculated by a minor modification of the approximate method of Sec. II for wall apertures, the details of which can be found in Ref. 23. We write

$$D(\omega)=\tfrac{1}{2}\pi bs^2[\Gamma'(\omega)-i\Delta'(\omega)], \tag{32}$$

where Γ' and $-\Delta'$ are nondimensional real and imaginary parts of D. The dependencies of Γ' and Δ'on Strouhal number fL/U are shown by the solid curves in Fig. 10(b) for a cavity of aspect ratio $b/L=1$. It is easily verified that energy is absorbed from the applied pressure gradient by the mean flow when $\Delta'>0$, i.e., for fL/U less than about 0.9. However, the motion is unstable, i.e., $D(\omega)$ is singular in Im $\omega>0$, since it may be verified that the computed values of $\Gamma'(\omega)$ and $\Delta'(\omega)$ do not satisfy the Kramers–Kronig formulas (7).

The behavior of $D(\omega)$ for complex ω can be estimated by using the relations (8) to derive an analytic approximation of the type (9) (in which κ must be replaced by $\omega s/U$, $s=1/2L$). The broken curves in Fig. 10(b) represent a four-pole analytic approximation with parameters

$$\alpha_1=-5.5,\quad \kappa_1=1.5,\quad \beta_1=4.5;\quad \alpha_2=-1.4,\quad \kappa_2=2, \tag{33}$$

$$\beta_2=-1.5;\quad \alpha_n=0,\quad n>2.$$

There are two poles in the upper half plane at $\omega s/U=\pm 2+1.5i$, which correspond to a mode that grows exponentially like $\exp\{1.5Ut/s\}$, and oscillates with Strouhal number

$fL/U = 0.64$. This value is close to measured Strouhal numbers for low Mach number flow over a shallow cavity in air,[5] and in water tunnel experiments for $b/L = 1$.[25]

V. CONCLUSION

Self-sustained oscillations of flow over an aperture or wall cavity are a consequence of shear layer instability. In this paper a linear perturbation theory of these oscillations is discussed for cases where the shear layer thickness is small and can be modeled by a vortex sheet. The motion in the neighborhood of the aperture has been assumed to be incompressible. For one-sided flow over a wall with circular or rectangular apertures, the frequencies of self-sustaining oscillations are identified with the real parts of the complex poles of the Rayleigh conductivity in the upper half of the frequency plane. A similar theory is applicable to shallow wall cavities, whose dimensions are very much smaller than the characteristic wavelength of the cavity tones, but the instabilities now correspond to poles of a drag coefficient.

In both cases we have estimated the lowest order (first ''stage'') Strouhal number of the unstable motion by approximate analytic continuation from the real frequency axis. For a shallow, rectangular wall cavity the predicted Strouhal number is about 0.64, which is in good agreement with available data in air at low Mach numbers. The minimum Strouhal number for flow over a wall aperture is predicted to be typically of order 0.65–0.8; it is weakly dependent on aperture aspect ratio, and decreases slowly with increasing wall thickness.

ACKNOWLEDGMENTS

This work was sponsored by the Office of Naval Research under Grant No. N00014-95-1-0318, administered by Dr. Patrick L. Purtell. It is a pleasure to acknowledge the benefit of discussions with Dr. William K. Blake during the preparation of the paper.

[1] D. Rockwell, ''Oscillations of impinging shear layers,'' AIAA J. **21**, 645–664 (1983).
[2] J. E. Rossiter, ''The effect of cavities on the buffeting of aircraft,'' Royal Aircraft Establishment Technical Memorandum 754 (1962).
[3] L. F. East, ''Aerodynamically induced resonance in rectangular cavities,'' J. Sound Vib. **3**, 277–287 (1966).
[4] N. M. Komerath, K. K. Ahuja, and F. W. Chambers, ''Prediction and measurement of flows over cavities—a survey,'' AIAA Paper 87-022 (1987).
[5] K. K. Ahuja and J. Mendoza, ''Effects of cavity dimensions, boundary layer, and temperature on cavity noise with emphasis on benchmark data to validate computational aeroacoustic codes,'' NASA Contractor Report: Final Report Contract NAS1-19061, Task 13 (1995).
[6] C. K. W. Tam and P. J. W. Block, ''On the tones and pressure oscillations induced by flow over rectangular cavities,'' J. Fluid Mech. **89**, 373–399 (1978).
[7] W. K. Blake and A. Powell, ''The development of contemporary views of flow-tone generation,'' in *Recent Advances in Aeroacoustics*, edited by A. Krothapali and C. A. Smith (Springer-Verlag, New York, 1986), pp. 247–325.
[8] J. C. Bruggeman, ''Flow induced pulsations in pipe systems,'' Ph.D. thesis, Eindhoven University of Technology, 1987.
[9] J. C. Bruggeman, A. Hirschberg, M. E. H. van Dongen, A. P. J. Wijnands, and J. Gorter, ''Flow induced pulsations in gas transport systems: analysis of the influence of closed side branches,'' J. Fluids Eng. **111**, 484–491 (1989).
[10] J. C. Bruggeman, A. Hirschberg, M. E. H. Van Dongen, and A. Wijnands, ''Self-sustained aero-acoustic pulsations in gas transport systems: experimental study of the influence of closed side branches,'' J. Sound Vib. **150**, 371–394 (1991).
[11] M. C. A. M. Peters, ''Aeroacoustic sources in internal flows,'' Ph.D. thesis, Eindhoven University of Technology, 1993.
[12] P. C. Kriesels, M. C. A. M. Peters, A. Hirschberg, A. P. J. Wijnands, A. Iafrati, G. Riccardi, R. Piva, and J. C. Bruggeman, ''High amplitude vortex induced pulsations in gas transport systems,'' J. Sound Vib. **184**, 343–368 (1995).
[13] Lord Rayleigh, *Theory of Sound* (Dover, New York, 1945), Vol. 2.
[14] M. S. Howe, M. I. Scott, and S. R. Sipcic, ''The influence of tangential mean flow on the Rayleigh conductivity of an aperture,'' Proc. R. Soc. London Ser. A **452**, 2303–2317 (1996).
[15] A. D. Pierce, *Acoustics, An Introduction to its Principles and Applications* (AIP, New York, 1989).
[16] E. C. Titchmarsh, *Introduction to the Theory of Fourier Integrals* (Clarendon, Oxford, 1967), 2nd ed.
[17] M. I. Scott, ''The Rayleigh conductivity of a circular aperture in the presence of a grazing flow,'' M.S. thesis, Boston University, 1994.
[18] D. G. Crighton, ''The Kutta condition in unsteady flow,'' Annu. Rev. Fluid Mech. **17**, 411–445 (1985).
[19] L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media* (Pergamon, Oxford, 1960).
[20] B. Noble, *1958 Methods based on the Wiener-Hopf Technique* (Pergamon, London, 1958; reprinted by Chelsea, New York, 1988).
[21] F. C. DeMetz and T. M. Farabee, ''Laminar and turbulent shear flow induced cavity resonances,'' AIAA Paper 77-1293 (1977).
[22] S. A. Elder, T. M. Farabee, and F. C. DeMetz, ''Mechanisms of flow-excited cavity oscillations at low Mach number,'' J. Acoust. Soc. Am. **72**, 532–549 (1982).
[23] M. S. Howe, ''Edge, cavity and aperture tones at very low Mach numbers,'' J. Fluid Mech. **330**, 61–84 (1997).
[24] G. K. Batchelor, *An Introduction to Fluid Dynamics* (Cambridge U. P., Cambridge, England, 1967).
[25] C. B. Burroughs and D. R. Stinebring, ''Cavity flow tones in water,'' J. Acoust. Soc. Am. **95**, 1256–1263 (1994).

Measurements of small-caliber ballistic shock waves in air

Roland Stoughton
Science Applications International Corporation, 10260 Campus Point Drive, San Diego, California 92121

(Received 10 June 1996; revised 14 April 1997; accepted 28 April 1997)

Megahertz bandwidth pressure measurements of airborne ballistic shock waves from bullets are presented which confirm weak-shock theoretical predictions of waveform shape, period, amplitude, and the scaling of these quantities with shock propagation distance. Wavefront distortions and amplitude variations are quantified versus shock propagation distances of 3–55 m and agree with predictions based on linear acoustic propagation through simple turbulence models. Observed rise times probably exhibit a mix of turbulence and molecular relaxation effects. © *1997 Acoustical Society of America.* [S0001-4966(97)03808-3]

PACS numbers: 43.28.Tc, 43.25.Cb [LCS]

INTRODUCTION

Ballistic shock waves from bullets present an interesting test of weak-shock theories and recently have been exploited, through digital signal processing, to locate snipers.[1,2] Theoretical modeling[3,4] of the amplitude and period of these shocks, and comparison of observations with the predicted scaling relations[4] were done in the 1940s. More recently, Bass *et al.*[5] have measured these shocks with an emphasis on their rise time. Rise times of these shocks,[5] as well as those of aircraft sonic booms,[6] are significantly larger then predicted from viscosity and still air. Atmospheric turbulence[7–9] and molecular relaxation[5,10,11] have been used to explain this excess, and both are established as important mechanisms. We repeat some of the earlier ballistic shock wave measurements, confirm quantitative agreement with the weak-shock predictions of shock wave period and amplitude, and also measure the effects of atmospheric turbulence on wavefront shape and peak amplitude. Good agreement is seen between our observations of wavefront spatial distortions and peak amplitude variations and the predictions of the theory of linear wave propagation through a turbulent medium.[12] This theory should be useful for predicting the conditions under which these wavefronts retain their integrity and hence when they can be exploited for practical tasks like projectile tracking and shooter localization. The linear wave theory also matches the observed rise times; however, this may be fortuitous since molecular relaxation probably is important, and since finite amplitude effects in the shock propagation may inhibit turbulent thickening of the shock leading edge.[6,10] Our intention is not to resolve the controversy over rise times. Rather it is to present high-quality confirmatory measurements of the ballistic shock waveform dependence on miss distance, and novel measurements of the spatial coherence and amplitude variations of these shocks.

In Sec. I we summarize the predictions of weak-shock theory for a still atmosphere. Observations are presented in Sec. II and compared to these predictions. Section III discusses the effects of atmospheric turbulence and compares our observations of wavefront distortions and rise times with linear propagation theory. A discussion is given in Sec. IV, and conclusions in Sec. V.

I. WEAK-SHOCK PREDICTIONS

Several projectile diameters and farther from the trajectory, the ballistic shock wave has attained a characteristic ''N'' shape, as in Fig. 1. Hence the name ''*N* wave'' which is often used. This limiting form arises as fine structure in the shock wave associated with the projectile shape is swept forward in time in the case of positive pressure perturbations, and back in the case of negative pressure perturbations, into only two discontinuities. The peak pressure amplitude $p_{\max}$ and period T of the N wave are related to projectile length l, diameter d, distance from the trajectory b, and Mach number M. By modifying the linear acoustics isentropic result with the finite-amplitude flow characteristic curves, Whitham[3] derives

$$p_{\max}=\frac{0.53P_0(M^2-1)^{1/8}}{b^{3/4}}\frac{d}{l^{1/4}}, \tag{1}$$

$$T=\frac{1.82Mb^{1/4}}{c(M^2-1)^{3/8}}\frac{d}{l^{1/4}}, \tag{2}$$

where P_0 is ambient pressure and c is sound speed.

Taking explicit account of energy dissipation, DuMond *et al.*[4] derive expressions which reduce to Eqs. (1) and (2) for large b/d, but show steeper dependencies of $P_{\max}$ and T on b for smaller b/d, where ''small'' is not determined. Their observations do indicate a steeper dependence for the larger caliber projectiles, consistent with the form of their predictions. Atmospheric conditions, instrumentation properties, and experimental procedure were not reported by DuMond *et al.*

Rise time of the leading and trailing edges of the N wave due to viscosity is predicted to be[13]

$$t_{\text{rise}}=\frac{\lambda}{c}\frac{P_0}{p_{\max}}, \tag{3}$$

where λ is the molecular mean free path, $\lambda\doteq 8\times10^{-8}$ m. For typical $p_{\max}/P_0>10^{-3}$ in the experiments this yields rise times in the submicrosecond range. When molecular relaxation effects are included,[5,11] the predicted t_{rise} comes out significantly larger than Eq. (3), but still in the microsecond range for small miss distances, leading to an experimental

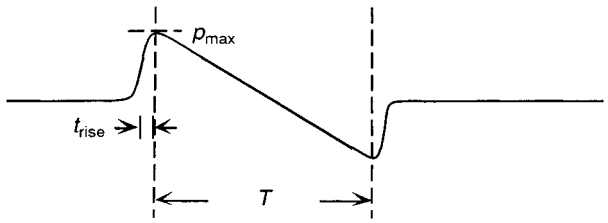


FIG. 1. Asymptotic shape of the ballistic shock wave.

requirement for piezoelectric transducers and megahertz sample rates if the shock waveform is to be resolved.

II. OBSERVATIONS OF BALLISTIC SHOCK WAVES

Data were collected in August 1995 at a local rifle range by firing past fixed condenser and piezoelectric microphones at known miss distances of 3–55 m. Data also were collected with condenser microphones at miss distances up to ~200 m at Camp Pendleton Marine Base, California, in October 1995. Although small-caliber ballistic shock waves were seen clearly out to shock propagation distances of ~300 m, quantitative analysis is limited primarily to the August data set, for which ground interaction effects were much better controlled, sound channel focusing was much less uncertain, and which was taken at a higher digital sampling rate.

In the August data set, pressure signals were burst-digitized with 12-bit precision at 2-MHz sample rate over 1-ms time windows. Experimental geometries were arranged so that all shocks arrived at near-normal incidence to the transducers' surfaces, fell within these time windows, which were triggered by a condenser microphone just ahead of the sensors to be recorded, and were free of ground interaction out to 20-m miss distance. The near-normal incidence angles were achieved through knowledge of muzzle velocity and ballistic drag coefficient, and confirmed by arrival times at our planar array. A horizontal linear arrangement of four piezoelectrics was used to measure shot-to-shot wavefront spatial distortions. Table I lists the sensor properties. The piezoelectric sensors were mounted flush with $3\times3\times1/4$ in. aluminum baffles. Figure 2 shows raw responses for four different sensor types. The long-discharge-time piezoelectric should most faithfully represent the early part of the incident waveform. Departure of the later part of the waveform from an ideal N wave at this small miss distance probably is due to the finite baffle size. Intrinsic rise times of these piezoelectrics are $<1\ \mu$s for normal incidence. The short-discharge piezoelectric response resembles the time derivative of the long-discharge piezoelectric, and is particularly appropriate for measuring leading edge arrival times. Both have polarity reversed in Fig. 2. The two condenser microphone responses exhibit ringing associated with the detailed dynamics of the membrane, cavity, and protective grid.

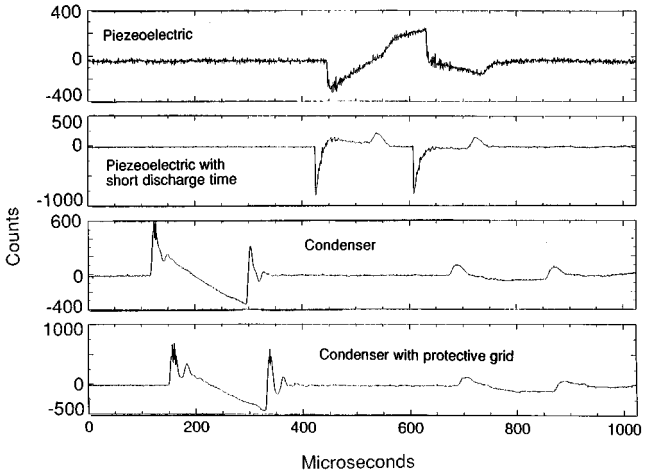


FIG. 2. Responses of different transducer types to the ballistic shock wave. 5.56-mm ammunition, 3-m miss distance. Data were digitized at 2-MHz sample rate. Transducers are described in Table I.

Figure 3 shows piezoelectric responses for four different miss distances from 3 to 55 m, and illustrates the behavior predicted in Eqs. (1) and (2). Also obvious is the lengthening of rise time and increasing distortion of the N wave with increasing miss distance.

For quantitative comparison with predictions of amplitude and period, the condenser microphone responses were used since a more accurate factory calibration was available for them. Period was measured as the time between the leading and trailing edges, and amplitude as the intercept of (i) a linear fit to the linear descending portion of the waveform, and (ii) the leading edge time. These quantities were insensitive to the details of the ringing in the condenser responses illustrated in Fig. 2. Observed shock wave amplitudes and periods are summarized in Figs. 4 and 5. Behavior is very close to the nominal relations, Eqs. (1) and (2). In evaluating (1) and (2), we must calculate the Mach number at the point on the trajectory where that part of the shock wave is emitted which will eventually hit the sensors. This was done by measuring this velocity with a chronograph for several rounds from the same ammunition lot. Measured velocity varied $<2\%$, and agreed with Remmington's ballistic data. Emission Mach was 2.7 for the 5.56-mm-diam and 1.9 for the 7.62-mm-diam projectiles, for the ambient sound speed of

TABLE I. Properties of the different transducers.

Sensor	ACO 7017	PCB 102A07	PCB 132A40	
Type	1/4 in. condenser	quartz piezoelectric	ceramic piezoelectric	
Dynamic range	50–164 dB	131–205 dB	115–171 dB	
Bandwidth	5–70 000 Hz	1–500 000 Hz	10 kHz–1 MHz	(Long discharge time)
			60 kHz–1 MHz	(Short discharge time)
Rise time				
Normal	~2 μs	2 μs	$<1\ \mu$s	
Oblique	25 μs	8 μs	3 μs	

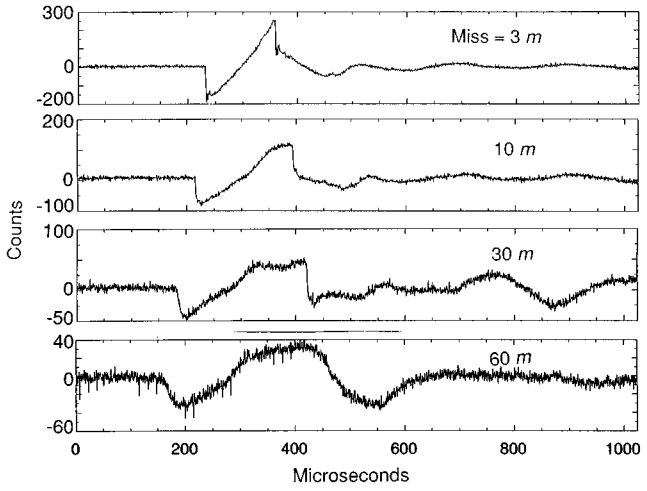


FIG. 3. Evolution of ballistic shock wave with miss distance. 5.56-mm ammunition, miss distances 3–55 m. Vertical scale (ADC counts) changes from plot frame to plot frame.

355 m/s (40 °C). In addition to a small bias, the amplitude behavior for the 5.56-mm projectiles shows a significant departure from the $b^{-3/4}$ power law at the larger miss distances. This could be due to sound channel focusing. The propagation path was 2 m off the ground on a hot day with solar heating. The sound-speed profile $c(z)$ was observed to be upward refracting and of curvature $\partial^2 c/\partial z^2$ in the sense that

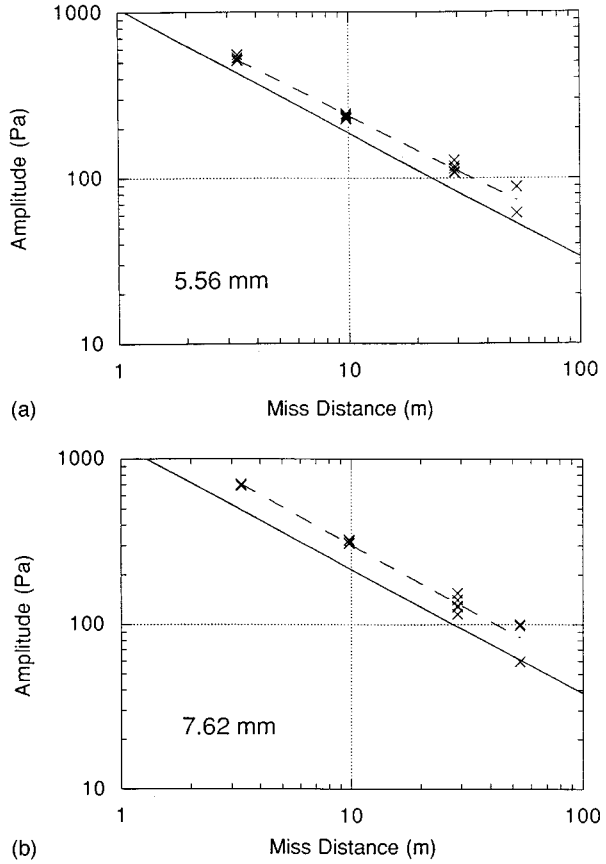


FIG. 4. Peak amplitude of N wave versus miss distance. (a) 5.56-mm ammunition, (b) 7.62-mm ammunition. Dashed lines are best fits to data. Solid lines are predictions from Eq. (1). Numbers of data points varied from 3 to 6 at any miss distance.

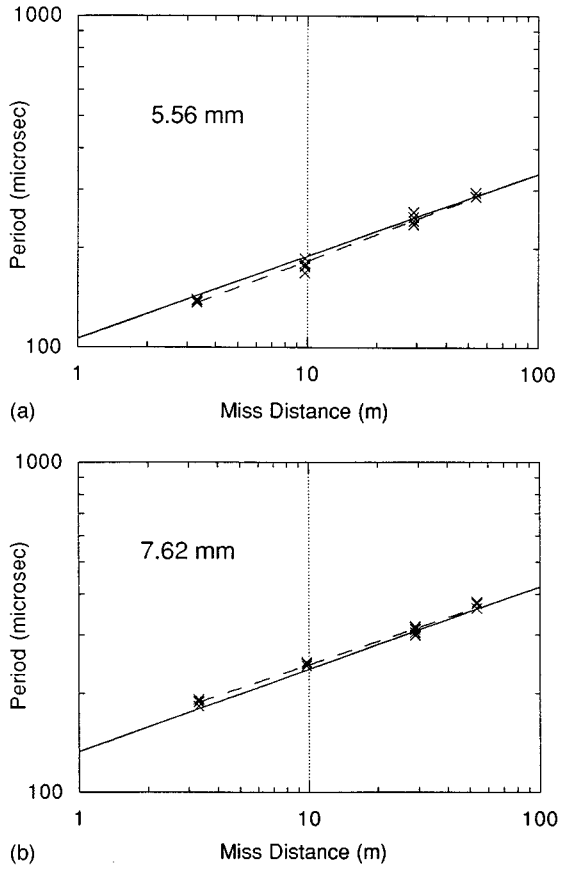


FIG. 5. Period of N wave versus miss distance. Dashed lines are best fits to data. Solid lines are predictions from Eq. (2).

would cause excess received amplitude, with this excess increasing with miss distance. The influence of $c(z)$ was not addressed by DuMond *et al.* Without accurate synoptic measurements of wind shear and temperature along the propagation path, we can only produce an order-of-magnitude estimate of the effect. Taking the observed temperature values of $\partial c/\partial z \doteq -0.3\ \mathrm{s}^{-1}$, $\partial^2 c/\partial z^2 \doteq 0.3\ \mathrm{m}^{-1}\,\mathrm{s}^{-1}$ gives an intensity focusing which reaches double the free-field value at $R = 100$ m, and the excess intensity scales as the square of propagation distance. Also, ground interaction becomes significant at the 30- and 55-m miss distances. We therefore distrust the data points at $R = 30$ m and $R = 55$ m when trying to make an accurate determination of the amplitude scaling with range. Using the points at $R = 3$ and 10 m, however, should keep the focusing effects at or below the 1% level, and should have no ground reflected paths in the data time window.

The effects of propagation uncertainties on period T are more complex. Again, we may take the 3- and 10-m points as most reliable. The resulting estimates for the power-law exponents are given in Table II. Determinations are consistent with the asymptotic rules $P_{\max} \sim b^{-3/4}$ and $T \sim b^{1/4}$ found by Whitham[3] and DuMond *et al.*[4]

Figure 4 shows agreement not only with the predicted scaling relations, as was shown by DuMond *et al.*,[4] but with the constant coefficients in Eq. (1), to within the $\pm 30\%$ experimental uncertainties in amplitude calibration. Figure 5 shows small systematic departures of measured periods from

TABLE II. Estimates of power-law exponents for $p_{max} \sim b^n$ and $T \sim b^m$. Data at 3- and 10-m miss distance only were used. Uncertainties represent 90% confidence intervals, based on scatter of repeated measurements at each miss distance.

Ammunition	5.56 mm	7.62 mm
n	-0.738 ± 0.027	-0.734 ± 0.015
m	0.230 ± 0.018	0.249 ± 0.012

Eq. (2) at 3- and 10-m miss distances. But these are in the opposite sense for the two projectile sizes. We have been unable to explain them as a measurement bias.

To compare measured rise times with Eq. (3) we use the long-discharge-time piezoelectric responses and define observed rise time to be the ratio of maximum pressure to maximum rate of increase of pressure, where this rate was estimated by least-squares line-fitting the portion of the leading edge between 10% and 90% of peak amplitude. This procedure was found to be unbiased with respect to, and gave smaller variance in repeated measurements than, the 10%–90% rule with smoothing. Figure 6 shows that the observed rise times are an order of magnitude longer than the weak-shock prediction, Eq. (3). As mentioned in Sec. I, a similar discrepancy was noted in the 1970s for aircraft sonic booms;[6,7] molecular relaxation and propagation through atmospheric turbulence both contribute in amounts that depend on the exact geometry and environmental conditions of the propagation.[6–9,14] Reference 5 attempts to match small-caliber ballistic shock rise times with molecular relaxation, obtaining order-of-magnitude agreement. However, their agreement is the worst for their indoor data set where turbulence is effectively absent. Deducing the physical mechanism is complicated by the roughly similar trend of rise time with propagation distance (decreasing peak pressure) predicted from molecular relaxation to that predicted from the accumulation over distance of turbulence effects. Further discussion is given in Sec. IV.

III. EFFECTS OF ATMOSPHERIC TURBULENCE

Our measurements support estimation of wavefront distortions down to the ~ 1-μs level, and amplitude variations to the $\sim 10\%$ level. Here we present these estimates and interpret them in terms of the effects of atmospheric temperature and velocity fluctuations, using a propagation theory appropriate to linear waves and stationary atmospheric fluctuation statistics. In our experiment, the short horizontal propagation path suggests a stationary model. Discussion of nonlinear propagation effects is deferred to Sec. IV. We also compare predictions of this theory to the observed rise times, although this comparison is complicated by both nonlinearity and molecular relaxation.

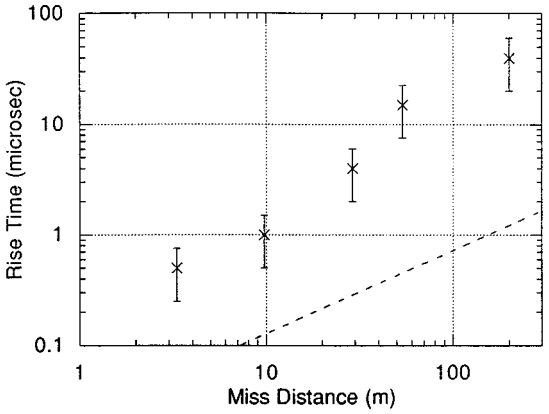


FIG. 6. Rise time of N wave versus miss distance, 5.56-mm ammunition. Dashed line is prediction from Eq. (3), which assumes a homogeneous atmosphere. The observation at 200-m miss distance comes from the October 1995 data set with condenser microphones and 100-kHz sampling.

Levels of turbulence were not measured during the data collection, but we can make predictions with an assumed spectrum using its level as a free parameter. We find that the observed wavefront distortions, amplitude variations, and lengthened rise times, and their dependence on shock propagation distance all are predicted by the same turbulence strength and spectral shape.

A. Turbulence model

We adopt a turbulence energy spectrum over isotropic wave number

$$F(k) \sim k^{-p}, \quad \lambda_0 < k^{-1} < L, \tag{4}$$

where λ_0, L are the inner and outer scales of the turbulence. L also is identified with the correlation length. Typical observed correlation behavior and fractional mean-squared sound-speed fluctuations μ^2 near the ground indicate[15,16]

$$L \sim 1 \text{ m}, \quad 10^{-7} < \mu^2 < 10^{-5}, \quad 5/3 < p < 2,$$

where $p = 5/3$ is the Kolmogorov spectrum and $p = 2$ arises when the turbulence is composed of random vortex sheets. For this range in the exponent p, results are not very sensitive to the inner scale λ_0.

B. Saturated and unsaturated regimes

The mechanics of converting a turbulence model to predictions of wavefront distortions are well established.[17] A very condensed discussion will be repeated here. The controlling parameters, usually denoted Φ and Λ are, respectively, the rms phase fluctuation due to sound-speed fluctuations integrated along a line from source to receiver, and a measure of the average number of independent fluctuations present in a Fresnel zone centered on that line. We have[17]

$$\Phi^2 = q_0^2 \mu^2 R \int_{-\infty}^{\infty} \rho(u)\, du, \tag{5}$$

$$\Lambda = R/6L^2 q_0, \tag{6}$$

where $q_0 = 2\pi f/c$, f is frequency, R is the propagation distance, and $\rho(u)$ is the correlation function, normalized so that $\rho(0) = 1$. For $p = 5/3$, Eq. (5) becomes

$$\Phi^2 \doteq 0.4 q_0^2 \mu^2 RL. \tag{7}$$

The most fundamental distinction in analyzing the propagation for any R, f, p is between ''unsaturated'' regimes, where $\Phi \Lambda^{p/4} < 1$ and geometrical optics is a good approximation, and ''saturated'' regimes, where $\Phi \Lambda^{p/4} > 1$ and significant amplitude fluctuations and waveform distortions

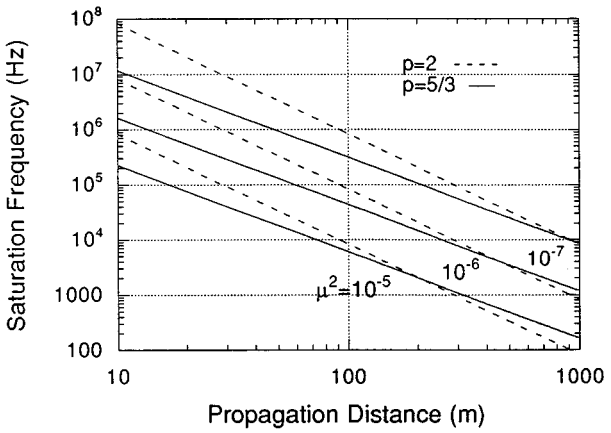


FIG. 7. Boundaries between saturated and unsaturated propagation, from Eqs. (8) and (9). Saturation occurs above the line. Solid lines are for turbulence spectral power-law index $p=5/3$, dashed lines are for $p=2$. $c=340$ m/s was used.

occur.[12] This distinction can be expressed as a saturation frequency, f_{sat} using (6) and (7),

$$p=5/3 \rightarrow \Phi^2 \Lambda^{5/6}=1 \rightarrow f_{\text{sat}} \doteq 1.3 c L^{4/7} \mu^{-12/7} R^{-11/7}, \quad (8)$$

$$p=2 \rightarrow \Phi^2 \Lambda=1 \rightarrow f_{\text{sat}} \doteq 2.5 c L \mu^{-2} R^{-2}. \quad (9)$$

This is illustrated in Fig. 7. Using $\mu^2=10^{-6}$ for example we see that at propagation distances >60 m the $\sim$MHz frequencies associated with the shock discontinuities will be saturated. The slight distortions of the entire waveform seen in Fig. 2 at the larger miss distances also are consistent with an approach to saturation for frequencies of ~ 10 kHz at ranges ~ 60 m indicated in Fig. 7.

C. Wavefront coherence

Because the dominant wavelengths in the shock waveform are clearly unsaturated at $R<60$ m, we may use geometrical optics approximations to compute arrival time variance and covariance between sensors. For separations transverse to the propagation direction we find arrival time covariance[17]

$$D_t^2 \equiv E[(\tau(x)-\tau(x+\Delta x))^2] = 2.2\mu^2 \frac{RL}{c^2}\left(\frac{\Delta x}{L}\right)^{5/3}, \quad \lambda_0 \ll \Delta x \ll L \quad (10)$$

$$= 0.8\mu^2 \frac{RL}{c^2}, \quad \Delta x \sim L, \quad (11)$$

where these have been derived for the case $p=5/3$.

These two precise limiting forms can be joined with the approximation

$$D_t^2 = 2\sigma_t^2(1-\rho_t(\Delta x)), \quad (12)$$

where

$$\sigma_t^2 = 0.4\mu^2 \frac{RL}{c^2}, \quad (13)$$

$$\rho_t(\Delta x) = (1+2.75(\Delta x/L)^{5/3})^{-1}. \quad (14)$$

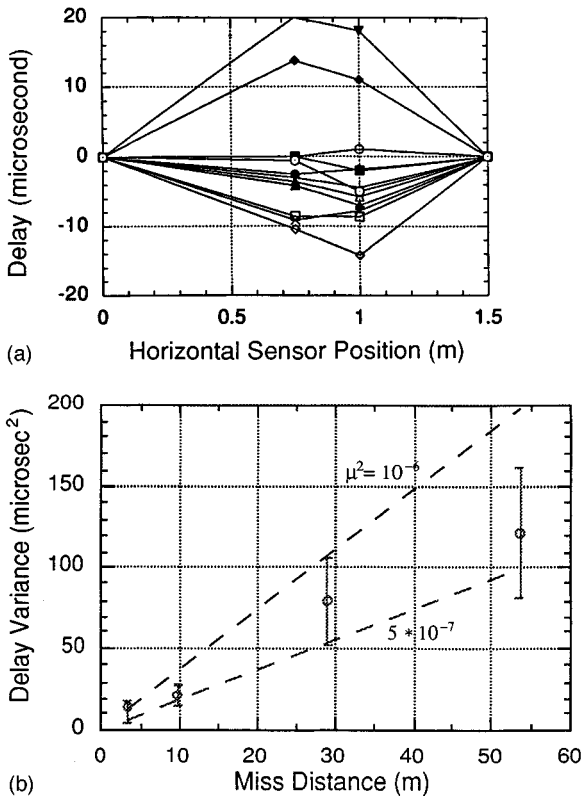


FIG. 8. (a) Residual shock wavefront distortions with respect to a plane wavefront passing through the two end-point sensors in a four-sensor horizontal array. 30-m miss distance. (b) Combined variance of the two middle sensors' residual arrival times versus miss distance. Dashed lines are predictions from Eq. (12), for two different assumed turbulence strengths.

σ_t^2 represents the total variance in acoustic travel time along a ray path from source to receiver; $\rho_t(\Delta x)$ is the correlation of this delay in the transverse direction.

Equations (10) and (11) were derived for a point source and receiver. Our shock wave is generated by a moving source and exists as an expanding conical wavefront with the projectile at the vertex. Coherences over separations in a plane containing the trajectory therefore may be better predicted based on a line source than a point source. This would affect the numerical coefficients in Eqs. (10), (11), (13), and (14). However, conclusions will be insensitive to this change since we will infer μ^2 from the data. Therefore the calculations in Ref. 17 have not been redone for a line source or moving point source.

Observations of horizontal coherence were done by firing repeated shots at the same nominal geometry about 2 s apart past a 1.5-m horizontal array of four short-decay piezoelectric sensors. The 2-s interval was chosen as somewhat longer than the turbulence decorrelation time for scales ~ 1 m in the light winds (1 to 2 m/s) present during the data collection. To compare observations with theory, the shot-to-shot variation in wavefront tilt had to be removed from both, since in the data this was contaminated with shot-to-shot variation in bullet velocity and trajectory. This was done by replacing the four arrival times with their residuals from a line through the two end-point sensor arrival times. The two end-point times then become zero [Fig. 8(a)]. The variances of the two nonzero times were then compared to predictions

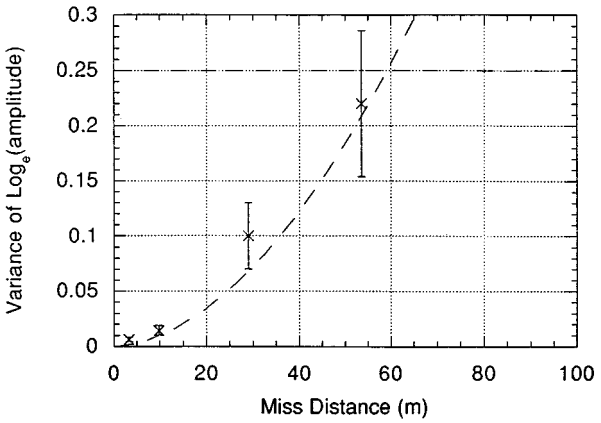


FIG. 9. Amplitude variations compared to predictions based on the turbulence spectral level $\mu^2=7\times10^{-7}$ inferred from Fig. 8. The saturation boundary is at a vertical coordinate of 0.25.

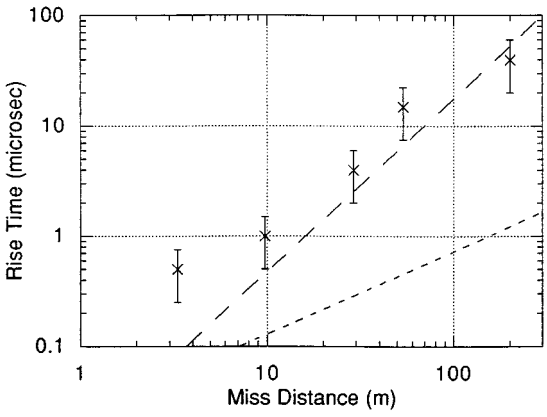


FIG. 10. Rise time versus predictions based on still air (short dash) and turbulent air (long dash) with turbulence level inferred from Fig. 8.

[Fig. 8(b)], where the predictions used Eqs. (12)–(14) to express these variances as a linear transformation of the covariance between the four sensors. The inferred level of μ^2 is 7×10^{-7}, consistent with other observations in similar conditions.[15] Our data were collected over sand in full sun during late-morning hours, in very light winds, with an ambient temperature of $\sim40\,^{\circ}$C.

D. Amplitude fluctuations

Using the spectral level inferred in Sec. III C from the arrival time variations we can predict the amplitude fluctuations and rise times. Variance of log (amplitude) is related to Λ, Φ as[17]

$$\sigma^2_{\log a}=\tfrac{1}{4}\Lambda^{p/2}\Phi^2 \tag{15}$$

and is equal to 1/4 at the boundary between saturated and unsaturated regimes. Using (5) and (6) we find

$$\sigma^2_{\log a}=0.022L^{-2/3}\mu^2R^{11/6}(2\pi f_0/c)^{7/6},\quad p=\tfrac{5}{3}, \tag{16}$$

which depends explicitly on frequency. By using data from the short-discharge-time piezoelectrics, we can arrange a well-defined characteristic frequency f_0 for the observed signals so that data may be compared with (16). The frequency response of these piezoelectrics is bounded below by the discharge time and above by the rise time. Defining f_0 as the centroid of the frequency response we are able to derive

$$f_0=t_d^{-1}\,\ln(t_d/t_r), \tag{17}$$

where t_d is the (exponential) discharge time and t_r the sensor response rise time. Examination of typical sensor responses to shocks at small miss distance gives $t_d=4\,\mu$s, $t_r=0.5\,\mu$s. Substituting Eq. (17) into Eq. (16) we compare with data as shown in Fig. 9, where amplitude has been measured as the peak in the response, and the curve represents Eq. (16). Evidently the predictions based on spectral level inferred from the wavefront coherence measurements are a good fit to the observed amplitude fluctuations.

E. Rise times

A rough prediction of shock waveform rise times due to turbulence can be made by arguing that all frequencies greater than the saturation frequency have their relative phases scrambled and cannot contribute to a sharp leading edge. This suggests

$$t_{\rm rise}\doteq f_{\rm sat}^{-1} \tag{18}$$

with $f_{\rm sat}$ from Eq. (8). More sophisticated calculations of linear-wave pulse spreading are possible,[18] but are not supported by our environmental data. The predictions using Eqs. (8) and (18) are plotted with the rise-time data in Fig. 10. Here again we set $\mu^2=7\times10^{-7}$ as inferred from the wavefront coherence measurements, and take $L=1$ m. At the shortest miss distance the observed rise time is set by the sensor response rise time of $\sim1/2\ \mu$s. The data point at the largest miss distance is from an October 1995 data set, in which propagation was across a valley, farther above the ground and therefore in weaker turbulence than the other data points which were from the August data set. The prediction therefore fits the data quite well where it should, and much better than the predictions for still air, Eq. (3).

IV. DISCUSSION

The consistency of the observed wavefront distortions and amplitude variations with turbulence predictions using only the one free parameter μ^2 suggests that the theory is appropriate for these observables and that the actual sound-speed variations had $\mu^2\sim7\times10^{-7}$. Changing the assumed spectral index p, and correlation length L within reasonable limits would have changed the inferred μ^2, but would not have affected the agreements between theory and data significantly.

The successful prediction of rise times may be fortuitous, since molecular vibrational relaxation effects are predicted to be significant, and since the wave is nonlinearly bound. The only data set comparable to ours which eliminates the effects of turbulence is that from ''range 4'' given in Ref. 5, which was an indoor range. Those rise-time data do not correlate well with the authors' predictions based on vibrational relaxation. The authors attribute this to an instru-

mental rise-time response limit of 2 μs. However, a limit of 2 μs does not explain the discrepancy between their ''range 4'' data and the theory. More recent modeling of rise time[11] coupling the nonlinearity with molecular relaxation shows longer rise times and large sensitivity to humidity (fraction of air made up of water molecules), which might be capable of explaining our observations without turbulent spreading.

As pointed out in Ref. 5, nonlinearity of the shock wave leads to resharpening on length scales which should be much shorter than the distances over which we are hypothesizing accumulation of turbulence effects. This argument would seem to invalidate our linear-wave prediction of rise times in turbulence. It does not undermine the turbulence explanation for the wavefront distortions and amplitude variations, since even finite-amplitude wavefronts are advanced, retarded, and focused by variations in local ambient sound speed. It is beyond the scope of this paper to try to resolve authoritatively the issue of the relative importance of turbulence and molecular relaxation in the rise times. In the context of aircraft sonic boom rise times, nonlinear[7] and nonstationary[6,9] propagation models have been employed, with order-of-magnitude success, and numerical prediction codes are available that show good agreement with each other and with at least some data sets.[14]

V. CONCLUSIONS

Measurements of small-caliber ballistic shock wave amplitude and period were made that confirm predictions based on weak-shock theory. Both the power-law exponents (Table II) and the constant coefficients agree closely with predictions. Despite the nonlinear nature of the shocks, the theory of linear acoustic wave propagation in turbulence is able to model successfully, with only one free parameter, the observed wavefront distortions, amplitude variations, and rise times of these shock waves, although the rise times would be more defensibly modeled with a combination of nonlinear wave propagation and molecular relaxational effects.

[1] G. L. Duckworth, D. C. Gilbert, and J. E. Barger, ''Acoustic counter-sniper system,'' Proc. SPIE **2938**, 262–275 (1996).
[2] R. Stoughton, ''SAIC SENTINEL acoustic counter-sniper system,'' Proc. SPIE **2938**, 276–284 (1996).
[3] G. B. Whitham, ''The flow pattern of a supersonic projectile,'' Commun. Pure Appl. Math. **V**, 301–348 (1952).
[4] J. W. M. DuMond, E. R. Cohen, W. K. H. Panofsky, and E. Deeds, ''A determination of the wave forms and laws of propagation and dissipation of ballistic shock waves,'' J. Acoust. Soc. Am. **18**, 97–118 (1946).
[5] H. E. Bass, B. A. Layton, and L. N. Bolen, ''Propagation of medium strength shock waves through the atmosphere,'' J. Acoust. Soc. Am. **82**, 306–310 (1987).
[6] A. Pierce, ''Statistical theory of atmospheric turbulence effects on sonic-boom rise times,'' J. Acoust. Soc. Am. **49**, 906–924 (1971).
[7] K. J. Plotkin and A. R. George, ''Propagation of weak shock waves through turbulence,'' J. Fluid Mech. **54**, 449–467 (1972).
[8] R. Raspet, H. E. Bass, L. Yao, and P. Boulanger, ''Statistical and numerical study of the relationship between turbulence and sonic boom characteristics,'' J. Acoust. Soc. Am. **96**, 3621–3626 (1994).
[9] P. Boulanger, R. Raspet, and H. E. Bass, ''Sonic boom propagation through a realistic turbulent atmosphere,'' J. Acoust. Soc. Am. **98**, 3412–3417 (1995).
[10] H. E. Bass and J. Ezell, ''Effect of vibrational relaxation on rise times of shock waves in the atmosphere,'' J. Acoust. Soc. Am. **74**, 1514–1517 (1983).
[11] A. D. Pierce, ''Atmospheric propagation of sonic booms,'' presentation at the 15th AIAA Aeroacoustics Conference, Long Beach, California, 1993 (unpublished).
[12] S. M. Flatté, ''Wave propagation through random media: Contributions from ocean acoustics,'' Proc. IEEE **71**, 1267–1294 (1983).
[13] Y. B. Zel'dovich and Y. P. Raizer, *Physics of Shock Waves and High-Temperature Hydrodynamic Phenomena* (Academic, New York, 1966), Vol. I, Sec. 23.
[14] J. P. Chambers, R. O. Cleveland, H. E. Bass, R. Raspet, D. T. Blackstock, and M. F. Hamilton, ''A comparison of computer codes for the propagation of sonic booms through realistic atmospheres utilizing actual acoustic signatures,'' NASA Conf. Publ. **3335**, 151–167 (1995).
[15] G. Daigle, J. Piercy, and T. Embleton, ''Effects of atmospheric turbulence on the interference of sound waves near a hard boundary,'' J. Acoust. Soc. Am. **64**, 622–630 (1978).
[16] H. E. Bass, L. N. Bolen, and J. Noble, ''Sound propagation through a turbulent atmosphere: experimental techniques and data analysis,'' The University of Mississippi report to U.S. Army Research Office, October 1987, DTIC # AD-A188891.
[17] *Sound Transmission Through a Fluctuation Ocean*, edited by S. M. Flatté (Cambridge U. P., Cambridge, England, 1979), Chap. 6.
[18] Reference 17, Chap. 8.

Further consideration of the waveguide propagation of ambient sound in the ocean-surface bubble layer

Zhen Ye[a)]
Institute of Ocean Sciences, Sidney, British Columbia, Canada and Department of Physics, National Central University, Taiwan, Republic of China

(Received 25 May 1995; revised 23 July 1996; accepted 25 January 1997)

Measurements of the ambient sound generated by breaking waves over the range 40–20 000 Hz reveal well-defined spectral peaks [D. M. Farmer and S. Vagle, J. Acoust. Soc. Am. **86**, 1897–1908 (1989)], suggesting the modulation of the ambient sound by the waveguide formed by the ocean-surface bubbles. The measurements show that the sound-speed profile in such a waveguide may be modeled by an exponential form. The theory of the waveguide propagation of the ambient sound is considered. It is shown that the sound propagation in the ocean-surface waveguides bears analogy with that in the optical waveguides formed by diffusion and can be solved analytically in a simple form. The cutoff frequencies of such waveguides are determined by a simple equation that incorporates only the waveguide parameters. Theoretical predictions of the spectral peaks, which seem to be associated with the modal cutoff frequencies, are compared favorably to the observations. Moreover, the present results are compared to the previous theory base on the model of inverse square sound-speed profiles in great details. It is shown that, although the two profiles are almost the same shape in the cases considered here, the sound propagation in the two waveguides is rather different. It is shown that the normal-mode propagation in the inverse square model is very sensitive to both source and receiver positions in addition to the waveguide parameters, and different modes can overlap, thereby making data fitting difficult. On the other hand, in the exponential model, the sound spectral function depends mainly on the waveguide parameters, e.g., sound-speed anomaly at the surface and the *e*-folding depth. Therefore the distinct spectral peaks observed are a stable prediction from the exponential waveguide model. The simplicity and the stability of the results can help solve the inverse problem, that is, ''given an observed distribution of spectral peaks, can the bubble distribution be inferred?'' © *1997 Acoustical Society of America.* [S0001-4966(97)00408-6]

PACS numbers: 43.30.Bp, 43.30.Nb, 43.30.Lz [JHM]

INTRODUCTION

Naturally generated ambient noise in the ocean may be created by breaking waves. The ambient sound can be used to infer the extent of the near-surface bubble layer through the observed masking of higher acoustic frequencies,[1] making it a useful probe for studies of many upper-ocean processes including gas transfer, ocean-surface turbulence, Langmuir circulation, and internal waves.[2–4]

Recently, elegant observations of ambient sound in the ocean-surface bubble layers have been reported by Farmer and Vagle.[5] The observations were obtained with a self-contained acoustic instrument that is deployed in the open ocean as a drifter suspended from a surface float by a rubber cord so as to decouple it from effects of wave motion. The instrument incorporates a broadband hydrophone to record the ambient sound, and a multifrequency inverted echo sounder with which to observe bubble clouds, allowing calculation of the resulting sound-speed profile. The observations were obtained in the FASINEX study approximately 200 miles SW of Bermuda and also moored in 140 m of water on the La Perouse Bank west of Vancouver Island, with the instrument at 24 and 14 m beneath the surface, respectively. The measurements of the ambient sound generated by breaking waves over the range of audio frequencies reveal well-defined spectral peaks, the frequency of which may remain generally consistent from one breaking event to the next, but which can change significantly over the course of a storm, or from one storm to another.

As discussed by Farmer and Vagle,[5] it seems unlikely that this previously unsuspected coherent fine structure in sound spectrum arises from coherent features in the sound-generation mechanism or fluid mechanical behavior of the breaking event; rather it is proposed that the explanation of these features lies in the modification of the sound by the near-surface bubble layer, i.e., normal-mode wave propagation. Thus a theory for waveguide propagation of the ambient sound has been proposed by Farmer and Vagle,[5] based on the concept of the trapping of a portion of the sound in the waveguide formed by the ocean-surface bubble layer.

The present paper is a continuation of our previous study[6] of the sound propagation in the waveguides formed by bubbles which are entrained into the ocean by breaking waves. The purposes of this paper are: (1) to further study the normal-mode propagation in the ocean surface; particularly we will consider propagation of individual modes; (2) to compare two similar waveguide models, namely, the exponential waveguide and the inverse-square waveguide; and

[a)]Electronic mail: zhen@joule.phy.ncu.edu.tw

(3) to compare results from the two waveguides with the observations. In this article, we examine the theory and show that the waveguide propagation of the ambient sound under the discussion bears analogy to the optical waveguides formed by diffusion which was first discussed by Conwell[7] and recently improved by the author.[8] When the sound-speed profile in the bubble layer is modeled as an exponential form, it has been shown that such a waveguide problem has a simple solution. We show that the number and positions of the well-defined spectral peaks can be determined by the waveguide properties through a simple equation, yielding favorable agreement with the experimental data. Moreover, in his comprehensive paper, Buckingham[9] has shown that the sound-speed profile for the ocean-surface bubble layer may also be fitted with an inverse-square form; and it has been argued that the theory based on the inverse-square sound-speed profile can also interpret the experimental data. Comparison is thus made between the two profiles in great detail. It is shown that, although the two profiles have some similarities in shape, the acoustic transmission in the two cases is rather different. It is argued that the inverse-square model may not be appropriate for explaining the well-defined spectral peaks observed in the experiments. Since the ocean is a complex system, different sound-speed profiles may be established under various conditions. Therefore the study of the waveguide propagation in this paper should not be considered as inclusive.

I. WAVEGUIDE PROPAGATION IN THE OCEAN-SURFACE BUBBLE LAYER

A. Cutoff frequencies

To simplify the problem, the presence of the surface waves is neglected and the ocean surface is assumed to be a flat pressure release surface; moreover, the range dependence is also ignored. These approximations are justified because a typical surface wavelength is of the order of 20–100 m, while the vertical e-folding scale of the bubble clouds is of the order of 2–3 m. Inclusion of surface waves and the situation right at the wave crest would form a complicated problem of rough waveguides[10] and will be the subject of future work.

Although some of the results below have been presented in Refs. 6 and 8, we purposely give some details for the reader's convenience. Moreover, through these details, the question of how to do inversion will have a clear answer. First we consider the eigenproblem. Under the above simplifications, the inhomogenous wave equation for a layered waveguide can be written in the cylindrical coordinates as[11]

$$\left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\,\frac{\partial}{\partial r}\right)+\frac{\partial^2}{\partial z^2}+\left(\frac{\omega}{c(z,\omega)}\right)^2\right]p(r,z,\omega)=0, \tag{1}$$

where r is horizontal range, z is the vertical coordinate, c is the sound speed, and ω is the radian frequency. For the normal-mode wave propagation in the waveguide, we can write the sound pressure field in the form[5,11]

$$p(r,z,\omega)=\sum_n R_n(r)F_n(z), \tag{2}$$

where the terms R_n refers to effects of cylindrical spreading and the eigenfunctions F_n satisfy

$$\left[\frac{d^2}{dz^2}+\left(\frac{\omega}{c(z,\omega)}\right)^2-k_{1n}^2\right]F_n(z)=0, \tag{3}$$

in which k_{1n} is a constant. The boundary conditions include a pressure release at the surface taken as the origin with z increasing upward, and that the solution must vanish at great depth:

$$F_n(0)=F(z\to-\infty)=0. \tag{4}$$

In principle, eigenequation (3) with arbitrary sound-speed profiles can only be solved numerically. Some numerical computation methods have been presented in a recent book by Jensen *et al.*[12] However, for certain simple profiles, exact solutions can be obtained, including the step-change, piecewise linear, inverse-square, and exponential profiles.[11,9,7,8] In the measurements of Farmer and Vagle,[5] it seems that on average the sound-speed profiles for the near-surface bubble layer have exponential nature. Representative measured profiles of the sound speed have been presented in Fig. 7 in Farmer and Vagle,[5] in which exponential curves have been shown to best fit the data. The resulting sound-speed profiles can be expressed in the form[5]

$$c(z)=c_1-\Delta c\,\exp(-|z|/d), \tag{5}$$

in which c_1 is the constant sound speed at great depth, Δc is the surface sound-speed anomaly, and d is the e-folding depth. Given the above sound-speed profile, the normal-mode wave propagation of the ambient noise is analogous to the optical wave propagation in waveguides formed by diffusion in semiconductor surfaces[7] and can be solved analytically.[8] It can be shown from statistical mechanics that the most probable depth distribution of bubbles is exponential; consequently the sound-speed profile is expected to be exponential as well because the sound speed depends on the bubble concentration linearly for low volume fractions.[5] This may support the observations of Farmer and Vagle.

Recognizing $\Delta c/c_1 \ll 1$, we can express the solutions to Eq. (3) with the exponential profile in terms of Bessel functions[7,8]

$$F_n(z)=AJ_{\nu_n}(\xi\,\exp(-|z|/d)), \tag{6}$$

where $J_{\nu_n}(\cdot)$ is the ν_nth order Bessel function of the first kind, and the parameters ν_n and ξ are defined as

$$\nu_n^2=4d^2(k_{1n}^2-k^2), \quad \text{with } k=\frac{\omega}{c_1}, \tag{7}$$

and

$$\xi=2kd(2\Delta c/c_1)^{1/2}. \tag{8}$$

The normal modes are determined by the eigenequation[5,7]

$$J_{\nu_n}(\xi)=0. \tag{9}$$

Here we note that, due to the boundary conditions, ν_n must be positive. Not recognizing this would lead to an overestimate of the number of the normal modes.

In Ref. 5, the authors solved eigenequation (9) by letting ν_n equal 0. The resulting equation is

$$J_0(\xi)=0, \tag{10}$$

which yields zeros at $\xi=2.405,\ 5.520,\ 8.654,...$, which the authors claim to correspond to the observed frequency bands (see Table II in Ref. 5). Obviously, these solutions are inadequate because the zeroth-order Bessel function of the first kind does not vanish as the argument approaches zero, and therefore cannot satisfy the correct boundary condition at great depth, i.e., Eq. (4).

The eigenequation in (9) can be encountered in many physical problems. To cite one example, we mention the quantum mechanical problem of the neutron–proton interaction.[13] The interaction potential can be approximated by an exponential function. Solving the Schrödinger equation for such potential would lead to an eigenequation which is exactly the same as Eq. (9). The solutions are called the bound states. Nuclear scientists try to infer the potential form from counting and analyzing these bound states.

The number of normal modes the waveguide will support is equal to the number of ν_n values for which $J_{\nu_n}(\xi)$ vanishes. This problem can be solved analytically by using the asymptotic expansion for the Bessel function[14]

$$J_\nu(\xi)\approx\sqrt{2/\pi}(\xi^2-\nu^2)^{-1/4}\cos[\sqrt{\xi^2-\nu^2}-\nu \times\cos^{-1}(\nu/\xi)-\pi/4]. \tag{11}$$

It was suggested in Ref. 8 that the asymptotic expansion in Eq. (11) can be used to obtain the eigenvalues for ν, i.e., $J_\nu(\xi)=0$. We find that the eigenvalues of ν_n are determined by

$$\sqrt{\xi^2-\nu_n^2}-\nu_n\cos^{-1}(\nu_n/\xi)-\pi/4=(n+1/2)\pi, \quad n=0,\pm1,\pm2,\ldots . \tag{12}$$

Further calculation leads to an elegant formula[8]

$$0\leqslant n\leqslant \xi/\pi-\tfrac{3}{4}, \tag{13}$$

from which we see that in order for the guide to support at least one normal mode, ξ must be great than $3\pi/4$. Since n must be an integer, we conclude from this relation that n can take any integral value in $[0,\xi/\pi-3/4)$, including zero. Therefore the number of modes (N) is given by

$$N=1+I, \tag{14}$$

where I is the integer formed by truncating the value resulting from $\xi/\pi-3/4$. Equation (13) shows that the minimum normal mode that the guide can support is $\xi_{\min}=3\pi/4$, and that the number of normal modes will increase by one when ξ increases by π. It is interesting to note that both $3\pi/4$ and π are two mathematical numbers that do not depend on the waveguide parameters Δc and d. This is a unique feature for an exponential waveguide. However, ξ does depend on the acoustic frequency and the waveguide parameters; and in the dispersive case these parameters are usually functions of the frequency.

Accordingly, the cutoff frequencies for the normal modes can be determined nicely by

TABLE I. Cutoff frequencies for the La Perouse and FASINEX examples.

Mode	La Perouse (Hz)	FV1 (Hz)	FASINEX (Hz)	FV2 (Hz)
1	3006	3000	1244	1314
2	7014	6132	2904	2695
3	11 022	9778	4563	4402
4	15 030	13 333	6222	5857
5	19 039	17 380	7881	7447
6	23 047	21 225	9540	9213
7			11 200	10 440
8			12 895	12 203
9			14 518	13 526
10			16 177	15 164
11			17 836	17 149
12			19 496	18 993

$$f_n=\left(\frac{1}{4\pi}\right)\xi_n\left(\frac{c_1}{d}\right)\left(\frac{c_1}{2\Delta c}\right)^{1/2} =\frac{c_1}{4d}\left(\frac{c_1}{2\Delta c}\right)^{1/2}\left(n-\frac{1}{4}\right), \quad n=1,2,3,\ldots . \tag{15}$$

Here we see that two adjacent bands are separated by a constant $c_1/4d(c_1/2\Delta c)^{1/2}$ for nondispersive waveguides. When waveguides are dispersive, the terms on the right-hand side of Eq. (15) will depend on the frequency, and the cutoff frequencies should be obtained self-consistently from the equation.

Consider applying the above results to the observations in the FASINEX and the La Perouse Bank studies by Farmer and Vagle.[5] The representative sound-speed profiles were presented in Fig. 7 of Farmer and Vagle. An interesting feature from the measurements is that the e-folding depth d is rather consistent parameter of approximately 1.4–1.5 m. This feature has also been mentioned in a review article by McDaniel.[15] It is also worth noting that the parameters Δc and d seem less dependent on the frequency at low frequencies (Refer to Fig. 7 in Ref. 5).

The results show that in the La Perouse case the surface sound-speed anomaly Δc and e-folding depth d vary only slightly with the frequencies and the sound-speed profile can be roughly fitted by

$$\Delta c=2.9\ \text{m/s}, \quad d=1.43\ \text{m}. \tag{16}$$

On the other hand, in the FASINEX case the waveguide parameters $(\Delta c,d)$ do depend on the frequencies rather significantly. For example, we have $\Delta c=14.3$ m/s, $d=1.56$ m for $f=40$ kHz, and $\Delta c=19.0$ m/s, $d=1.42$ m at 5 kHz. This shows significant dispersive features in the FASINEX case. At this stage, it is impossible to study the dispersive features in full scale because of the lack of data. In this article, we simply take the assumption of Ref. 5 that the waveguide is nondispersive and the sound-speed profile can be fitted by

$$\Delta c=19.0\ \text{m/s}, \quad d=1.5\ \text{m}, \tag{17}$$

for FASINEX. From Fig. 7 in Ref. 5, it seems that the parameters Δc and d are less sensitive to the frequency at low frequencies, thus implying that the theory is more accurate at the low frequencies.

In Table I, we present the theoretical prediction of the cutoff frequencies from Eq. (15). Here we use c_1

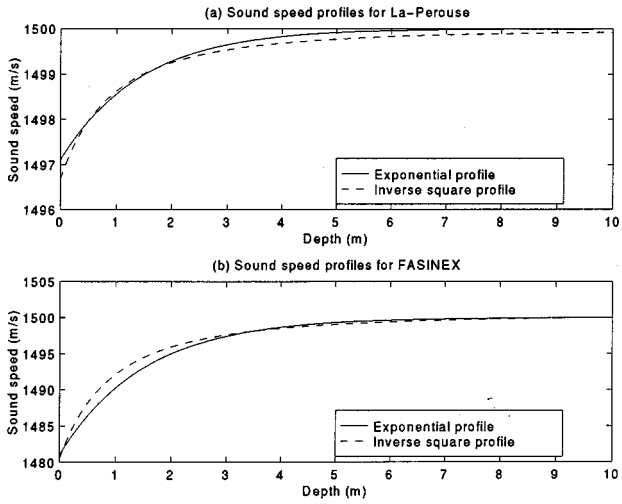


FIG. 1. Sound-speed profiles for (a) La Perouse and (b) FASINEX from exponential and inverse-square waveguide models.

$=1450$ m/s for La Perouse and $c_1=1500$ m/s for FASINEX. In the table, FV1 and FV2 are, respectively, for La Perouse and FASINEX from Table III of Farmer and Vagle.[5]

Inspecting the results in Table I and comparing them with the measurements shown in Fig. 10 of Farmer of Vagle,[5] we see that the cutoff frequencies may correspond to the spectral peaks in the observed sound spectra from individual breaking events. The agreement is remarkable, considering the simple model employed. But we notice some disagreements at high frequencies. The reasons for the discrepancy may lie in the ignored dispersive effects and rough-surface effects. Although the cutoff frequencies can correspond to the measured spectral peaks to certain extent, the theory would be incomplete without examining sound spectra.

Buckingham[9] has proposed an inverse-square sound-speed profile model to explain the observed data. The sound-speed profile is written as follows:

$$c(z)=c_0(1+z_1^2/(z+z_s)^2)^{-1/2}, \tag{18}$$

where the two parameters z_1, z_s are determined by fitting the experimental data. They are as follows:[9] $z_1=0.12$ m, $z_s=1.8$ m, and $c_0=1500$ m/s (La Perouse); $z_1=0.297$ m, z_s zs$=1.8$ m, and $c_0=1500.5$ m/s (FASINEX). Figure 1 shows the comparison between the exponential and the inverse-square sound-speed profiles. Here we see that the two are in fact very close to each other. The difference between the two is less than 0.03%. However, as will be shown in the following, the sound propagation in the two cases can be rather different.

B. Sound spectra and comparison with the existing theory

In the last section, the eigenproblem of wave propagation in an exponential waveguide was solved. In this section, we shall consider sound propagation in the presence of acoustic sources.

The sound generation in the ocean-surface bubble layer is a difficult problem. It is not the purpose of this paper to address the sound generation in the bubble layer; rather we assume a simplest acoustic source and investigate how the sound should propagate. Since the surface waveguide may also be modeled as an inverse-square form, as shown by Buckingham,[9] it is therefore rather imperative to compare the two sound-speed profiles with reference to the experiments. These will be the focus of this subsection.

Suppose an arbitrary acoustic source, $\rho(\mathbf{r},t)$, the wave equation can be written as

$$(\nabla^2+k^2)p(\mathbf{r})=-4\pi\rho(\mathbf{r},\omega), \tag{19}$$

where $k=\omega/c(z)$ and

$$\rho(\mathbf{r},\omega)=\int dt\ e^{i\omega t}\rho(\mathbf{r},t).$$

A formal solution to the above wave equation can be obtained with aid from Green's function which satisfies the following equation:

$$(\nabla^2+k^2)G(\mathbf{r},\mathbf{r}')=-4\pi\delta(\mathbf{r}-\mathbf{r}'), \tag{20}$$

and the boundary conditions stated earlier. Using Green's theorem, the solution to Eq. (19) can be written as

$$p(\mathbf{r})=\int d^3r'\ G(\mathbf{r},\mathbf{r}')\rho(\mathbf{r}',\omega). \tag{21}$$

The sound power spectrum, defined as $\langle|p|^2\rangle$, can be calculated from the following spatial correlation function:

$$D(\mathbf{r},\mathbf{r}')\equiv\langle p(\mathbf{r})p^{\dagger}(\mathbf{r}')\rangle = \int d^3r_1\ d^3r_2\ G(\mathbf{r},\mathbf{r}_1)G^{\dagger}(\mathbf{r}',\mathbf{r}_2)\langle\rho(\mathbf{r}_1,\omega)\rho(\mathbf{r}_2,\omega)\rangle \tag{22}$$

at $\mathbf{r}=\mathbf{r}'$, where $\langle\cdot\rangle$ refers to an ensemble average, and $\dagger$ for the complex conjugate operation. The source correlation function $\langle\rho(\mathbf{r}_1,\omega)\rho(\mathbf{r}_2,\omega)\rangle$ is unknown. If we assume uncorrelated random media, the source correlation function can be written as

$$\langle\rho(\mathbf{r}_1,\omega)\rho(\mathbf{r}_2,\omega)\rangle=S(\mathbf{r}_1,\omega)\delta(\mathbf{r}_1-\mathbf{r}_2). \tag{23}$$

Substituting this relation into Eq. (22), we obtain

$$D(\mathbf{r},\mathbf{r}')=\int d^3r'\ S(\mathbf{r}',\omega)|G(\mathbf{r},\mathbf{r}')|^2. \tag{24}$$

From the above derivation, it is clear that the acoustic source, represented by $\rho(\vec{r},t)$, can play a rather important role in defining the acoustic features of sound propagation in the ocean-surface bubbly layer. Unfortunately, this source function is often unknown, although a few models have been proposed for sound generation in the bubble layer.[15] However, we recognize that Green's function is another important quantity that can provide insight to the wave transmission in the bubble layer, as discussed in Ref. 9; it represents the intrinsic response of a system to a point source. Therefore, to make the problem manageable in this paper, it may be sufficient to investigate Green's function and see how the acoustic wave propagates in response to a point source, following the line of Ref. 9. The moral is that once Green's function is obtained and the sound source function is known, the actual sound spectrum can be calculated. In this paper,

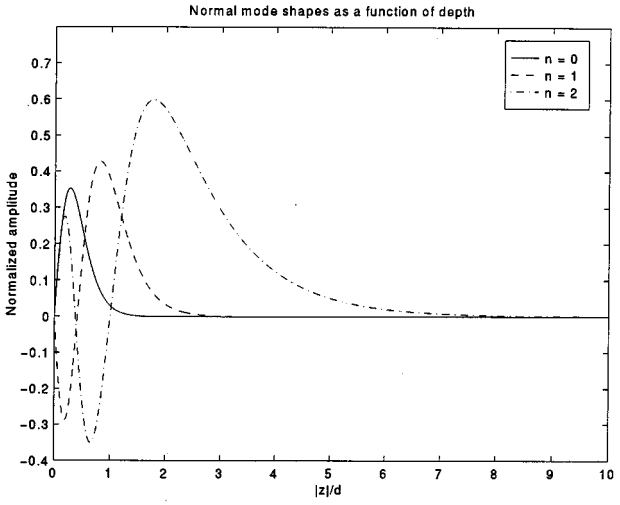


FIG. 2. Normal-mode shapes as a function of the normalized depth. Here $\xi=10$ has been assumed, which yields three normal modes.

we will consider Green's function. A study of sound propagation in the presence of a spatially distributed source in the bubble layer will be the subject of a forthcoming paper.

The Green's function of the wave equation involving the exponential profile can be solved following the standard Fourier–Bessel transformation method.[6,12] For convenience, we work in the cylindrical coordinates. For a point source at $(0,z_0)$, the solution can be obtained as follows:

$$G(z,r)=\frac{1}{2}\int_{-\infty}^{\infty} dq\; H_0^{(1)}(qr)Q(z,q), \tag{25}$$

with $H_0^{(1)}(qr)$ being the zeroth-order Hankel function of the first kind and

$$Q(z,q)=(-i)(2\pi d)\,\frac{J_v(t_>)}{J_v(\xi)} \times [H_v^{(1)}(\xi)J_v(t_<)-J_v(\xi)H_v^{(1)}(t_<)],$$

where $t_{\langle,\rangle}$ refers to the smaller or larger values of t and t_0, respectively. Here

$$t=\xi\,\exp(-|z|/(2d)), \quad t_0=\xi\,\exp(-|z_0|/(2d)),$$

and

$$q^2=k^2+\frac{v^2}{4d^2}.$$

A detailed derivation of the above result can be found in Ref. 6.

Using the well-known EJP cut,[12] the integral in Eq. (25) can be decomposed into a branch line integral and a normal-mode summation, resulting from the residues on the complex q plane. The branch line integral represents the contributions from spectral components being evanescent in range (see, e.g., page 125 in Ref. 12). Since the recording hydrophones are located outside the bubble layer, the contribution from the branch line integral may be neglected. Therefore the resulting Green's function can be approximated as

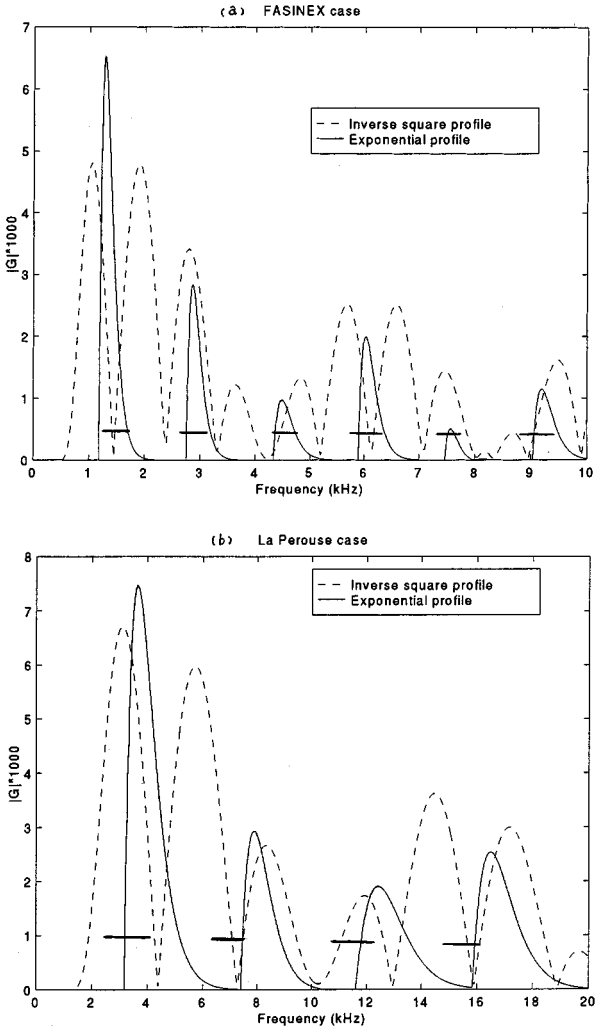


FIG. 3. Sound spectrum as a function of frequency for (a) FASINEX and (b) La Perouse. The horizontal stripes are experimental data from Ref. 5.

$$G(r,z)\approx\sum_{n=0}^{N}\left(\frac{\pi^2}{2d}\right)\frac{v_n J_{v_n}(t_0)J_{v_n}(t)H_{v_n}^{(1)}(\xi)}{[\partial J_v(\xi)/\partial v]|_{v=v_n}}H_0^{(1)}(q_n r), \tag{26}$$

where v_n is the solution of $J_{v_n}(\xi)=0$, and N has been given before. We note that in the exponential waveguide case, the normal-mode index starts from $n=0$, differing from the inverse-square waveguide.

For the sake of comparison, we now write down the solution to the inverse-square sound-speed profile as follows [see Eq. (31c) in Ref. 9]:

$$G_{\rm inv}(r,z)=-\frac{\pi}{8}\,Q\sqrt{(z+z_s)(z_0+z_s)}\sum_{m=1}^{\infty} \times\frac{\eta_m H_\mu^{(2)}(\eta_m z_s)H_\mu^{(1)}(\eta_m(z+z_s))H_\mu^{(1)}(\eta_m(z_0+z_s))}{\partial H_\mu^{(1)}(\eta z_s)/\partial\eta|_{\eta=\eta_m}} \times H_0^{(1)}(\sqrt{k^2-\eta_m^2}\,r), \tag{27}$$

where

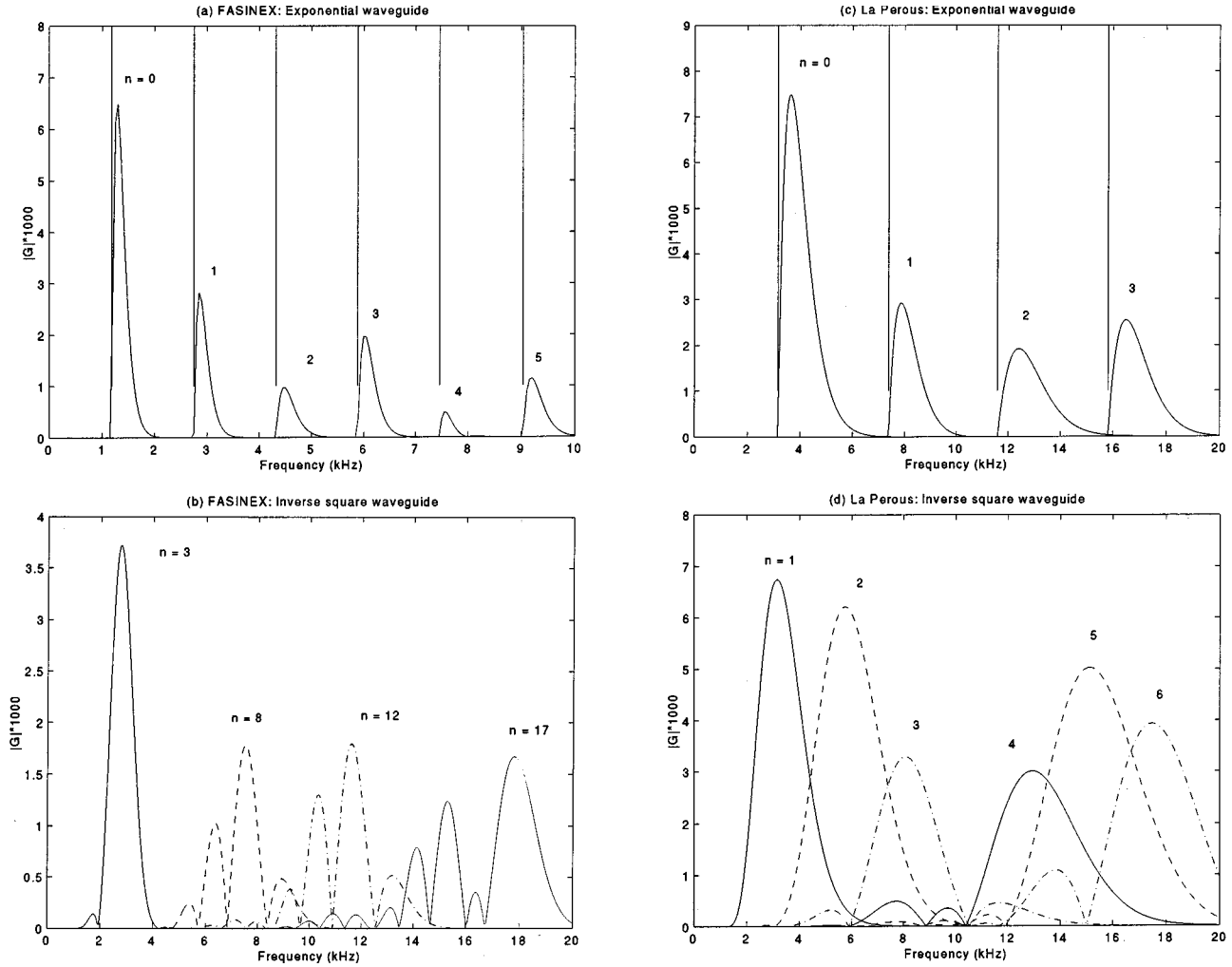


FIG. 4. Magnitude of the normal modes in the theoretical sound spectrum as a function of frequency: (a) FASINEX: exponential waveguide model; (b) FASINEX: inverse-square waveguide model; (c) La Perouse: exponential waveguide model; and (d) La Perouse: inverse-square waveguide model.

$$\mu = \sqrt{1/4 - k^2 z_1^2},$$

and the roots of the characteristic equation

$$H_\mu^{(1)}(\eta z_s) = 0$$

give the infinite eigenvalues η_m for modes $m = 1,2,3,\ldots,\infty$. Here we take Q as 4π to comply with the current definition of a point source (i.e., the result in Ref. 9 is multiplied by a factor of 4π.) Here we see that a striking difference between the exponential waveguide and the inverse-square waveguide is that, at a certain frequency, the exponential waveguide can only support a finite number of normal modes, while the inverse square can support an infinite number of normal modes, although not all these normal modes are important.

First let us look at the normal-mode shapes as a function of depth in the exponential waveguide. From the last subsection, the normal-mode solution is proportional to $J_{v_n}(\xi \exp(-|z|/d))$. As an example, we consider $\xi = 10$. In this case, there are three normal modes that the waveguide can support: $v_n = 6.09,\ 3.19,\ 0.89$. In Fig. 2, the normal-mode shapes have been plotted as a function of normalized depth $(|z|/d)$. Here we see that the normal modes decay rather quickly as the depth increases, and higher modes can penetrate deeper than the lower modes. This figure also shows that the energy is not necessarily only confined within the e-folding depth, i.e., $|z|/d = 1$.

Now we consider Green's function for the exponential waveguide with comparison to that for the inverse-square waveguide. Unless otherwise notified, the following parameters are used for numerical computation: source depth $z_0 = 1.5$ m, horizontal range $r = 10$ m, and sound speed at large depth $c = 1500$ m/s;[9] and the receiver depth at 14 m for the La Perouse and 24 m for the FASINEX.[5] Here a note should be made: Since in this subsection we emphasis on the comparison between the exponential and inverse-square profiles, we have take $c = 1500$ m/s even for the La Perouse case, following Ref. 9. This differs from $c = 1450$ m/s, the number we have used earlier. In Fig. 3, we plot Green's function as a function of frequency for the FASINEX and La Perouse cases. In the diagrams, the solid lines refer to the exponential waveguide results and the dashed lines refer to the inverse-square waveguide results, respectively. Here we see that in the exponential profile case, well-separated bands appear for both FASINEX and La Perouse cases; while in the inverse-

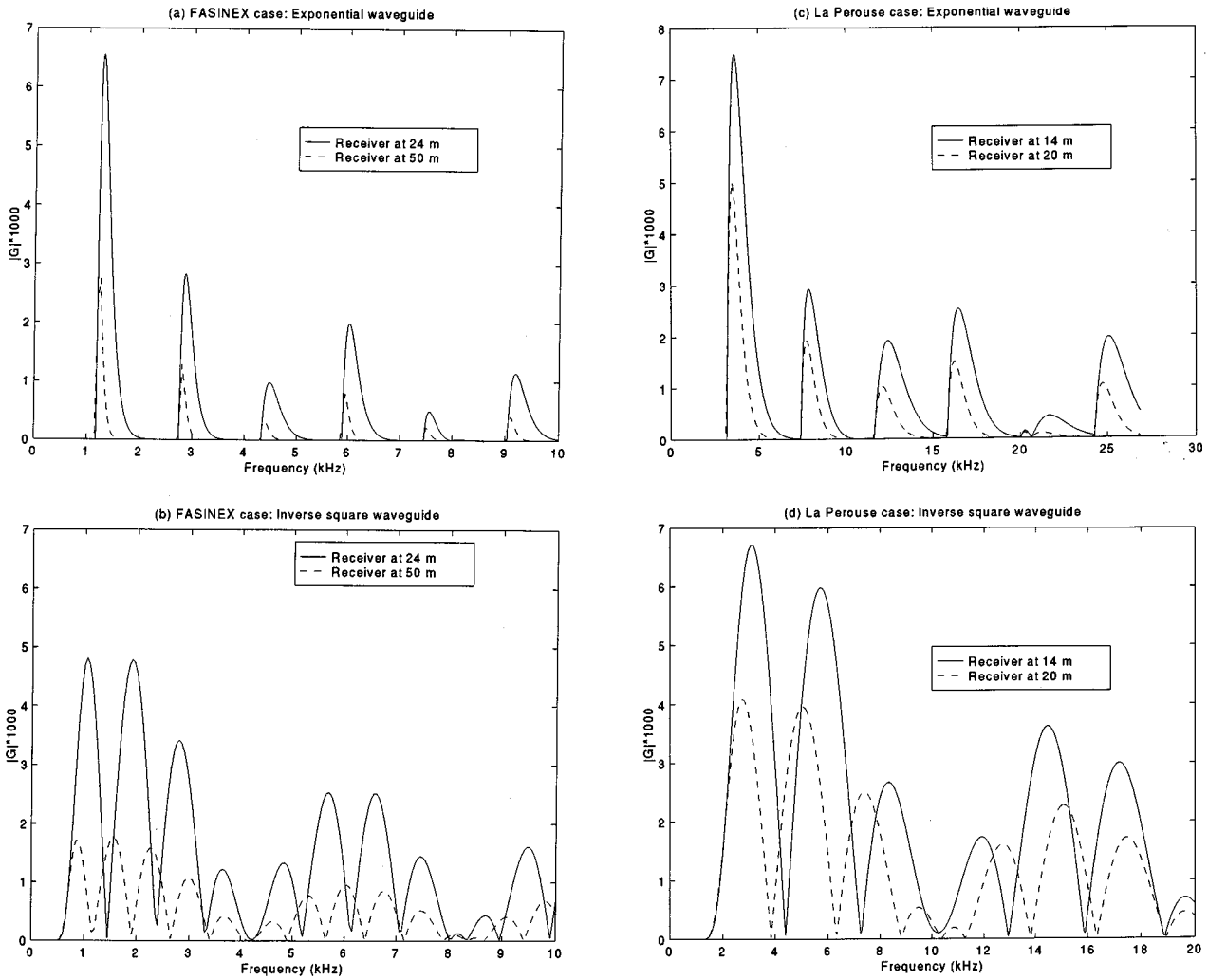


FIG. 5. Sound spectrum as a function of frequency for two receiver depths: (a) FASINEX: exponential waveguide model; (b) FASINEX: inverse-square waveguide model; (c) La Perouse: exponential waveguide model; and (d) La Perouse: inverse-square waveguide model.

square guide, more peaks appear. For example, in the FASINEX case, 10 peaks appear below 10 kHz in contrast to six peaks in the exponential waveguide; and seven peaks below 20 kHz in the inverse-square profile, compared to four peaks in the exponential waveguide for the La Perouse case. However, in both waveguides, the Green's functions have the same order of magnitude. We compare the results in Fig. 3 with the experimental results shown in color in Fig. 16 of Farmer and Vagle.[5] The reader is suggested to refer to the original data in Farmer and Vagle. The experimental data are presented as the horizontal stripes in the present figure. The first spectral data for La Perouse and the second spectral data for FASINEX are chosen for the comparison. It appears to the author that some well-defined isolated frequency bands do appear and seem to correspond to what the exponential waveguide theory predicts. Although the peak positions in the La Perouse case are not exactly at the positions predicted by the exponential theory due to the choice of waveguide parameters (the reason has been given earlier), the separation between the bands is remarkably consistent with the prediction of the exponential theory, in contrast to what the inverse-square model predicts. In Ref. 5, three breaking events are reported for each location, respectively. Here we only consider one event from each location. The inverse-square waveguide theory also predicts some well-defined spectral peaks for the La Perouse case, for a particular source depth. According to the inverse-square theory, the appearance of well-defined isolated spectral maxima in the La Perouse is fortuitous in the inverse-square model. As will be shown later, the overlapping of different normal modes in the inverse-square model may smear out the band structures. In the FASINEX case, the inverse-square theory predicts more peaks, and some spectral peaks are so closely spaced that they have been depicted as horizontal stripes in Ref. 9 to fit the observed spectral maxima.

We can investigate the contribution from each normal modes. Diagrams in Fig. 4 show the normal-mode contributions. Here we see clearly that in the exponential waveguide, each well-separated spectral band is actually given by individual normal modes, and there is no significant interference between normal modes. Only in the La Perouse does the $n=2$ mode overlap very slightly with $n=3$ mode at the tail [Fig. 4(c)]. The vertical lines in Fig. 4(a) and (c) refer to the cutoff frequencies derived using the theory developed in the

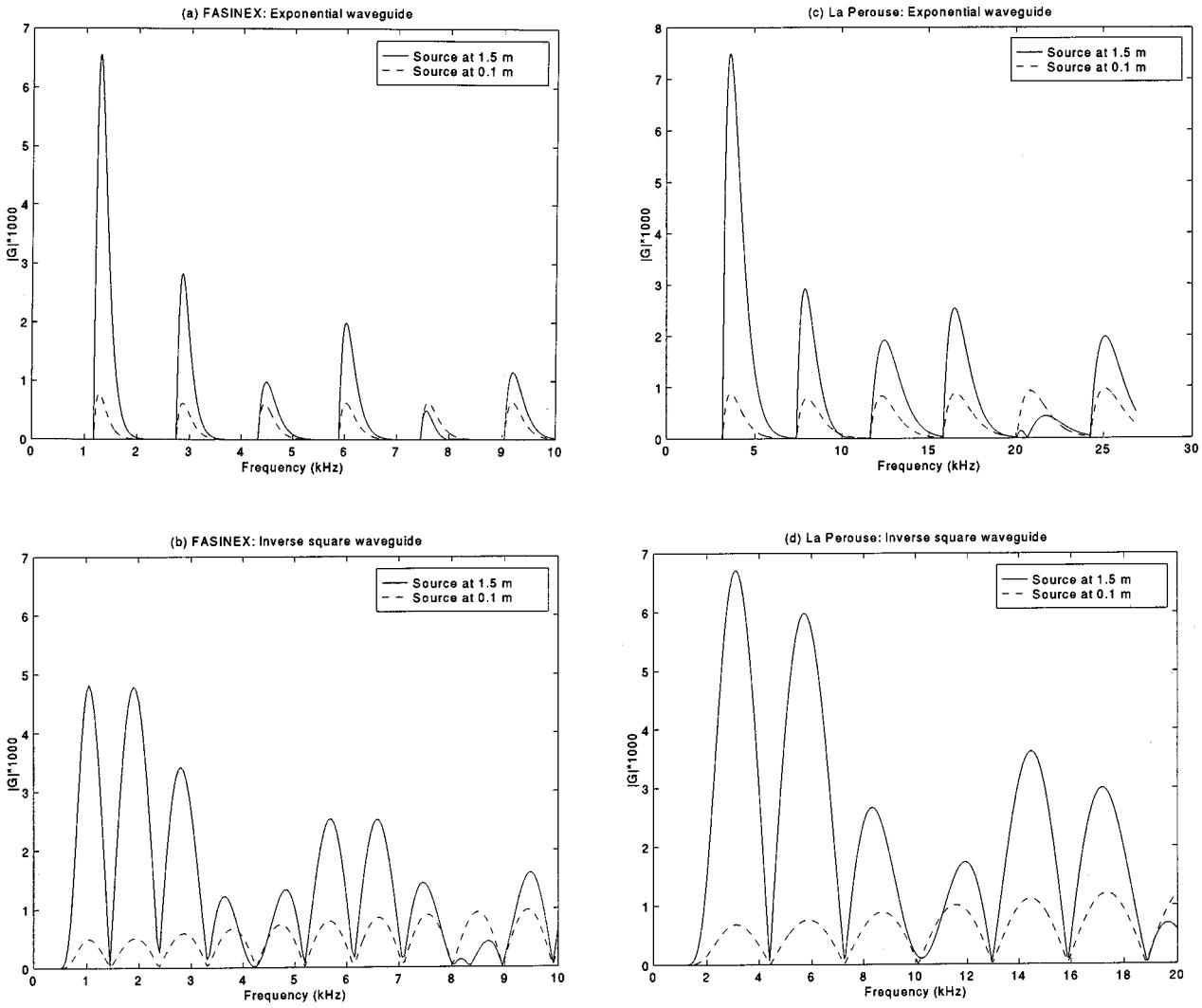


FIG. 6. Sound spectrum as a function of frequency for two source depths: (a) FASINEX: exponential waveguide model; (b) FASINEX: inverse-square waveguide model; (c) La Perouse: exponential waveguide model; and (d) La Perouse: inverse-square waveguide model.

last subsection. Here we also see that in the exponential waveguide, although the cutoff frequencies are not at the spectral peak positions, they are positioned only slightly toward the lower end from the peak positions, and the displacements are nearly identical. This supports the cutoff frequency theory, i.e., the separation of frequency bands can be approximately estimated from the cutoff frequencies. It is the author's opinion that the separation between bands is a more meaningful parameter rather than the absolute position of each band. The case with the inverse-square waveguide is rather different: The interference between different normal modes is obvious and significant. Such interference is much more pronounced in the FASINEX case. The final result is a coherent summation of several normal modes. In addition, each mode shows a multilobe structure, with more peaks appearing in the higher-order modes, adding complexity to the problem.[9]

We can also study the sensitivity of the sound spectrum with regard to the receiver depth. Figure 5 shows the sound spectrum as a function of frequency for two receiver depths. Here we see that in the exponential waveguide case, the spectral bands seem insensitive to the receiver depth except that the spectral peak magnitudes decrease with receiver depth, as expected. Here we note that a side band appears at the fifth band in the La Perouse [Fig. 5(c)]. In the inverse-square waveguide, things are different: The spectral structures can be altered significantly with changing receiver depth. More peaks appear at greater depth. However, it is recognized by the author that according to the inverse-square model, in spite of the additional bands, some bands may be grouped and regarded as forming broad stripes; and these stripes may be regarded as following the same spectral pattern for the different depths. For example, in the FASINEX case [Fig. 5(b)], the first three peaks (bands) for the receiver at 25 m may be regarded as forming a strip; the stripe may look similar to the stripe which could consist of the first four peaks for the case when the receiver is at 50 m. But, in Fig. 5(b) and (d), we do see the changes of the fine structures of the spectra. Since the exponential waveguide and the inverse-square waveguide are so different in spectral structure, changes with regard to receiver depths, it may be worth-

while to perform further experiments to determine which model is more plausible.

In Fig. 6, we inspect the sound spectrum as a function of frequency for two source depths. Again, in the exponential waveguide, the spectral structures almost remain the same for the two source depths. However, for the case in which the source is closer to the surface, the well-isolated peaks nearly have the same magnitude. In this case, because the source is farther away from the receiver, the magnitude is lower than when the source is at greater depth, except at the fifth band. In addition, the side band which appears at the fifth band in the La Perouse case disappears when the source moves closer to the surface [Fig. 6(c)]. For the inverse-square waveguide, at smaller source depths, the shapes of the sound spectra also more or less remain the same. The magnitude of each band is also almost the same, although a slight increase with increasing frequency is noticeable. In this case, it may be rather difficult or arbitrary to group the spectral bands into different stripes as to represent the observed bands.

From the above discussion, it is clear that the spectral structures (shapes) of the sound spectrum in the exponential waveguide seem to more rely on the properties of the waveguide; and in the inverse-square waveguide they are sensitive to both the waveguide parameters, and the receiver or source depth. When an extended source is considered, the exponential waveguide is expected to retain the well-defined isolated spectral structures, while in the inverse-square waveguide these fine structures may be smeared out; therefore the well-defined spectral peaks may not be predicted by the inverse-square model. However, in the author's opinion, the two theories are complimentary. In light of the complexity of oceanic processes, the ocean-surface bubble layer may sometimes be better modeled with the exponential waveguide, when the well-defined peaks are observed, and other times may be better modeled as the inverse-square waveguide; the worst case is that it can hardly be modeled as any well-defined waveguides.

II. CONCLUDING REMARKS

In this article, we have reconsidered the waveguide propagation of ambient noise in the ocean-surface bubble layer. The foregoing analysis presents an explanation of the observed peaks in the sound spectrum generated by breaking waves. It was found that the observed spectral peaks from individual wave breaking events may be explained in terms of the normal-mode wave propagation. It is shown that the cutoff frequencies of the normal modes for such ocean-surface waveguides can be written in a very simple formula, i.e., Eq. (15), in which the waveguide parameters appear in a simple form. The results show that two adjacent acoustic peaks may be separated by a constant value for nondispersion waveguides; or more generally, that a parameter ξ can be defined in terms of the waveguide parameters and two adjacent values of ξ are separated by a constant π. This is an important feature for exponential waveguides because once the surface sound-speed anomaly is measured, the e-folding depth of the bubble concentration may be easily inferred by measuring the separation of the two lowest spectral peaks, because the waveguides is less dispersive at low frequencies. This helps solve the inverse problem put forth by Farmer and Vagle.[5]

The exponential waveguide theory has been compared to the existing inverse-square theory. It is shown that although the two theories can model similar sound-speed profiles, distinctive differences between the two remain. In the exponential waveguide: (1) only a finite number of normal modes can be supported at certain frequency in contrast to the infinite number of normal modes in the inverse-square waveguide; (2) sound spectrum shows the well-defined and isolated frequency bands, in contrast to the interference between different normal modes in the inverse-square model; and (3) these well-defined structures are not sensitive to the source depth nor the receiver depth; in the inverse-square waveguide theory, the sound spectrum depends significantly on the source depth and the receiver depth. Further computation shows that the spectral structure is also not sensitive to the horizontal range of the source (not shown here). It may not be possible to infer the position of acoustic sources using the exponential waveguide theory.

Finally, it is worth noting that, as pointed out by Buckingham,[9] the spectral peaks may be smeared out together, due to ocean fluctuations, in which case observations of individual maxima could be the exception rather than the rule. For these complicated situations, the exponential waveguide model fails.

ACKNOWLEDGMENTS

D. M. Farmer and S. Vagle are thanked for useful discussions. Some results in the revised version are motivated by a discussion with M. J. Buckingham, not suggesting whether he agrees or not to the contents of this paper. In fact, Buckingham has raised a few concerns with the comparison between the exponential and inverse-square waveguides. He believes that the inverse-square model works better for the observed data. Grace Kamitakahara-King is thanked for useful suggestions and editorial corrections. The useful comments and constructive suggestions from the referees are also appreciated. The work received support from the Department of Fisheries and Oceans, Canada and the College of Science, the National Central University.

[1] D. M. Farmer and D. D. Lemon, ''The influence of bubbles on ambient noise in the ocean at high wind speeds,'' J. Phys. Oceanogr. **14**, 855–863 (1984).
[2] S. A. Thorpe, ''On the clouds of bubbles formed by breaking wind waves in deep water and their role in air-sea gas transfer,'' Philos. Trans. R. Soc. London, Ser. A **304**, 155–210 (1982).
[3] S. A. Thorpe, ''A model of the turbulent diffusion of bubbles below the sea surface,'' J. Phys. Oceanogr. **14**, 841–854 (1984).
[4] S. A. Thorpe, ''The effect of Langmuir circulation on the distribution of submerged bubbles caused by breaking wind waves,'' J. Fluid Mech. **142**, 151–170 (1984).
[5] D. M. Farmer and S. Vagle, ''Waveguide propagation of ambient sound in the ocean-surface bubble layer,'' J. Acoust. Soc. Am. **86**, 1897–1908 (1989).
[6] Z. Ye, ''On acoustic propagation in exponential ocean surface waveguides,'' J. Appl. Phys. **78**, 6389–6396 (1995).
[7] E. M. Conwell, ''Modes in optical waveguides formed by diffusion,'' Appl. Phys. Lett. **23**, 328–329 (1973).

[8] Z. Ye, ''Modes in optical waveguides formed by diffusion revisited,'' Appl. Phys. Lett. **65**, 3173–3175 (1994).
[9] M. J. Buckingham, ''On acoustic transmission in ocean-surface waveguides,'' Philos. Trans. R. Soc. London, Ser. A **335**, 513–555 (1991).
[10] F. G. Bass and I. M. Fuks, *Wave Scattering from Statistically Rough Surfaces* (Pergamon, Oxford, 1979).
[11] L. M. Brekhovskikh, *Waves in Layered Media* (Academic, New York, 1980).
[12] F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics* (AIP, New York, 1994).
[13] S. Flugg, problem 75 in *Practical Quantum Mechanics* (Springer-Verlag, New York, 1974).
[14] N. Bleistein and R. A. Handelsman, *Asymptotic Expansions of Integrals* (Holt, Rinehart and Winston, New York, 1975).
[15] S. T. McDaniel, ''Sea surface reverberation: A review,'' J. Acoust. Soc. Am. **94**, 1905–1922 (1993).

Sound scattering by an air bubble near a plane sea surface

Zhen Ye
Institute of Ocean Sciences, Sidney, British Columbia, Canada

Christopher Feuillade
Naval Research Laboratory, Stennis Space Center, Mississippi 39529

(Received 6 November 1996; accepted for publication 26 January 1997)

This paper presents a study of low-frequency sound scattering by an air bubble near a flat pressure-release surface. Using a self-consistent approach to describe multiple-scattering interactions, and a monopole approximation for the individual scatterer, the complete scattering amplitude for a bubble near a surface is derived. Radiation, thermal, and viscous damping effects are incorporated. The method leads to simple theoretical expressions which show that the presence of the surface modifies the resonance frequency and damping of the bubble, and modulates the scattering amplitude. The analytic formula for the modified resonance frequency compares favorably with earlier theoretical results, and with experimental data reported by M. Strasberg [J. Acoust. Soc. Am. **25**, 536–537 (1953)]. At off-resonance frequencies, when the bubble is close to the surface and thermal and viscous dampings are not included, the results obtained agree with those of the modal series solution presented by Gaunaurd and Huang [IEEE J. Ocean. Eng. **20**, 285–295 (1995)]. The two methods disagree at resonance frequencies, however, since the present approach predicts that the value of the scattering amplitude should decrease as the bubble is moved closer to the surface. This feature remains even when thermal and viscous dampings are not included. When the bubble is moved further away from the surface, oscillations appear in the scattering amplitude, in full accordance with the Lloyd's mirror effect. [S0001-4966(97)06407-2]

PACS numbers: 43.30.Bp, 43.30.Nb, 43.30.Lz [JHM]

INTRODUCTION

Sound scattering by air bubbles near the sea surface has become a very important issue in a number of current research applications, including the modeling of scattering from surface dwelling fish, and the understanding of oceanic fluxes and ambient noise in the upper ocean. The acoustic properties of an air bubble near a boundary were first studied by Strasberg,[1] who used an electrical capacitance method to calculate the effects of both rigid and free (pressure-release) surfaces on the pulsation frequency of a spherical bubble. The study of sound scattering by a spherical air bubble near a boundary was further advanced by Oguz and Prosperetti,[2] who calculated the linear oscillation frequency of a bubble in the vicinity of a slightly distorted plane free surface by a perturbation method, and Tolstoy,[3] who considered sound scattering by an air bubble near a thin elastic plate.

More recently, Gaunaurd and Huang[4] presented a benchmark modal series solution for the scattering problem of an air bubble adjacent to a pressure-release plane surface. Their result is particularly significant, because it is valid for a wide range of scattering angles, frequencies, and distances between the bubble and the boundary. The solution presented by these authors is also exact, so that it may be used to verify alternative approximate solutions. However, it has the disadvantage of being expressed in terms of complicated coupling coefficients, which are the solutions of an infinite, coupled, linear, complex equation, and can be found in terms of products of the Wigner-3j symbols.[4] For this reason, the actual computation based on the modal series solution may be quite tedious, as acknowledged by the two authors.

In a recent comment on the paper of Gaunaurd and Huang, Strasberg[5] noted a peculiar feature of their solution, namely that the peak values of the bubble scattering resonance function are shown to increase as the bubble is moved closer to the free surface. This result seems contrary to what one should intuitively expect. Strasberg attributes the apparent paradox to the significant reduction in radiation damping when other sources of damping are not included. In response to Strasberg, Gaunaurd and Huang[6] expressed their conviction that, by including the corrective effect of thermal and viscous damping, the unexpected trend would be reversed. However, their exact method of solution may not allow an easy inclusion of any mechanism limiting the magnitude of bubble pulsations at resonance except radiation damping. The work reported in this present paper stemmed from our motivation to consider this issue further, by including thermal and viscous damping, as well as radiation damping, in the problem.

It is instructive to note that although the modal series solution obtained by Gaunaurd and Huang is valid for a wide range of frequencies, their consideration was restricted to the important low-frequency regime, where the acoustic wavelength is considerably larger than the bubble radius. In this low-frequency limit, scattering by an air bubble is typically dominated by the radially pulsating ''monopole'' mode, even when a boundary is nearby. The assumption that this is the case has provided the basis for many studies. (See, e.g., Ref. 3.) After careful study, we believe that the problem of sound scattering by an air bubble near a free surface may be studied to an adequately high level of accuracy by an approximate method which utilizes this assumption, at least in the low-frequency regime. This regime is of particular inter-

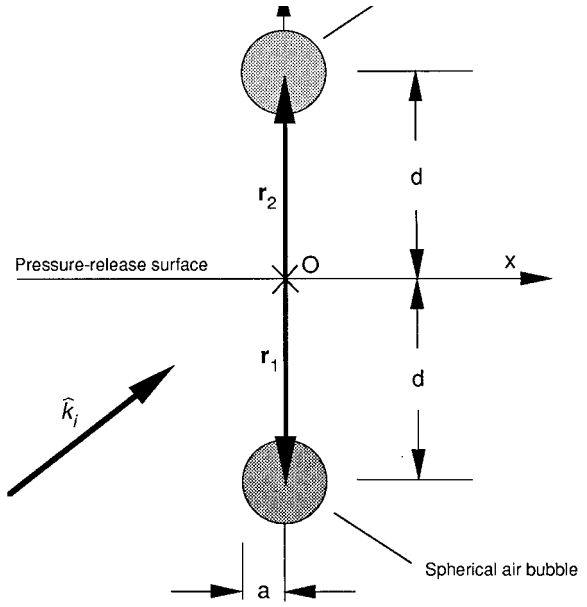


FIG. 1. Schematic diagram showing an air bubble near a flat pressure-release surface.

est, because it is here that the strong monopole resonance behavior of bubbles occurs, by which acoustic scattering is greatly enhanced. This has important implications for fish detection, and for the use of bubble population estimates to infer upper ocean dynamics.

This paper presents a study of low-frequency sound scattering by an air bubble near a flat pressure-release surface, based on the general method of images, and using the self-consistent method for treating the multiple scattering of acoustic waves pioneered by Twersky and others.[7,8] This self-consistent scheme was also used by Tolstoy to study superresonance phenomena in systems consisting of bubble doublets or triplets.[3] An alternative self-consistent method was used by Feuillade[9] to investigate the mechanism of super-resonances. In this paper, the sound scattering amplitude will be derived, incorporating radiation, thermal, and viscous damping. The formalism will be applied to several examples, and the results compared to those obtained by previous workers. In particular, we compare our results to those obtained from Strasberg's analytic solution,[1] and the exact modal series solution,[4] in order to validate the present approach. A major advantage of the method presented in this paper is its simplicity, and the clear insight it gives into the physical nature of this intriguing problem.

I. THE ANALYSIS

The problem considered in this paper is depicted schematically in Fig. 1. A spherical air bubble of radius a is located at a depth d beneath a plane, pressure-release surface. For convenience, the rectangular coordinate system is set up in such a way that the surface is in the x-y plane. The z axis is drawn vertically upward, and is shown passing through the center of the bubble (and also through its image reflected in the plane, of which more will be said later). As conventionally assumed, a continuous plane wave is considered to be incident on this bubble from below. The direction of incidence is described by the unit vector $\hat{k}_i$. The incoming plane wave is both scattered by the bubble and reflected from the surface, as shown. Subsequently, the wave field scattered from the bubble is reflected from the surface; while the surface reflected part of the incident wave, in turn, undergoes scattering from the bubble. This process is repeated many times to establish an infinitely recursive pattern of rescattering and rereflection between the bubble and the surface, causing the scattering characteristics of the bubble to change. The multiple-scattering process which ensues can be conveniently treated in a self-consistent manner.

First consider the air bubble located at $\mathbf{r}_1$. The scattered wave at $\mathbf{r}$ in response to an arbitrary incident wave $p_i(\mathbf{r})$ can be generally represented as[8]

$$p_s(\mathbf{r}) = \hat{f}(\hat{k}_i, \hat{k}_s) p_i(\mathbf{r}_1) \frac{e^{ik|\mathbf{r}-\mathbf{r}_1|}}{|\mathbf{r}-\mathbf{r}_1|}. \tag{1}$$

Here $\hat{f}(\hat{k}_i, \hat{k}_s)$ is a scattering amplitude operator dependent on the incident and scattering directions, where $\hat{k}_i$ is defined above, and $\hat{k}_s (= (\mathbf{r}-\mathbf{r}_1)/|\mathbf{r}-\mathbf{r}_1|)$ denotes the unit vector for the scattering direction. The propagation wave number is denoted by $k = \omega/c$, where ω is the acoustic angular frequency, and c is the sound speed in the liquid. It has been shown in many studies (e.g., Ref. 3) that, in the low-frequency domain (i.e., $ka \ll 1$), the scattering function of an arbitrary scatterer may be represented by the superposition of a radiating monopole and a dipole source. The latter, however, is usually negligible in the resonance scattering regime. Indeed, as shown by Twersky,[7] the scattering amplitude operator can be usually written as

$$\hat{f}(\hat{k}_i, \hat{k}_s, \mathbf{r}) = \left[1 + \frac{i}{2kr} \hat{D} + \left(\frac{i}{2kr} \right)^2 \frac{\hat{D}(\hat{D}-2)}{2} + \cdots + \left(\frac{i}{2r} \right)^n \times \frac{\hat{D}(\hat{D}-1\cdot 2)(\hat{D}-2\cdot 3)\cdots(\hat{D}-(n-1)n)}{n!} + \cdots \right] f(\hat{k}_i, \hat{k}_s), \tag{2}$$

where $f(\hat{k}_i, \hat{k}_s)$ is the customary scattering amplitude at far field, and

$$\hat{D} = \frac{-1}{\sin^2 \theta_s} [\partial^2_{\phi_s} + \sin \theta_s \partial_{\theta_s} (\sin \theta_s \partial_{\theta_s})],$$

in which θ_s, ϕ_s refer to the two polar coordinates for the scattering direction $\hat{k}_s$. In the present case, the scattering amplitude is dominated by the radially pulsating (i.e., ''monopole'') mode of the bubble, and can be considered isotropic as long as ka is smaller than about 0.35. This has recently been shown to be true even for deformed bubbles.[10] It is worth noting that all of the cases considered by Gaunaurd and Huang[4] fall well within this limit. In this frequency regime, the scattering amplitude can be approximated by [Eq. (A6.1.20) in Ref. 11]

$$\hat{f}(\hat{k}_i, \hat{k}_s, \mathbf{r}) \approx f = \frac{a}{\omega_0^2/\omega^2 - 1 - i\delta}, \tag{3}$$

where ω_0 is the resonance frequency given by

$$\omega_0 = \frac{1}{a}\sqrt{\frac{3\gamma P_0}{\rho_l}}. \tag{4}$$

Here P_0 is the hydrostatic pressure at the bubble's location, γ is the specific heat ratio (taken as ≈ 1.4), and ρ_l is the mass density of the surrounding liquid. The quantity δ is the frequency-dependent damping factor, a sum of the radiation (δ_r), thermal (δ_t), and viscous (δ_v) damping factors, i.e.,

$$\delta = \delta_r + \delta_t + \delta_v . \tag{5}$$

The radiation damping factor $\delta_r = ka$. The exact mathematical forms and typical values of the other damping factors for air bubbles in water are given in Appendix 6 of Ref. 11. These forms are used later in the numerical examples discussed in this paper.

Now let us consider the effect of the pressure-release boundary, shown in Fig. 1, on the wave scattered from the bubble. According to Brekhovskikh,[12] the effect of the boundary can be represented by introducing an image bubble with opposite source strength, symmetrically positioned above the surface. Therefore, writing the scattered field of the bubble generically as

$$p_s(\mathbf{r}) = A\,\frac{e^{ik|\mathbf{r}-\mathbf{r}_1|}}{|\mathbf{r}-\mathbf{r}_1|}, \tag{6}$$

where the constant A is to be determined, then the scattered field of the image bubble is simply

$$p_s'(\mathbf{r}) = -A\,\frac{e^{ik|\mathbf{r}-\mathbf{r}_2|}}{|\mathbf{r}-\mathbf{r}_2|}, \tag{7}$$

where $\mathbf{r}_2$ denotes the position of the image bubble. Again, in writing the above equations, all contributions from oscillation modes other than the pulsation mode have been ignored.

To determine A we shall employ an elegant and general self-consistent method, which has previously been used successfully by Tolstoy and Twersky to study multiple scattering processes.[3,7] The brief description of this method which follows, it must be stressed, is not intended to be rigorous, but is aimed to convey the basic principles that underlie the self-consistent approach. Consider a generic system of N scatterers. The scattered wave from the ith scatterer can be represented quite generally as

$$p_s^i = A_i\,\frac{e^{ik|\mathbf{r}-\mathbf{r}_i|}}{|\mathbf{r}-\mathbf{r}_i|} \quad (i = 1,2,\ldots,N) \tag{8}$$

and can clearly be written also in terms of the scattering function for this single scatterer as

$$p_s^i = \hat{f}_i\, p_{\mathrm{inc}}^i\,\frac{e^{ik|\mathbf{r}-\mathbf{r}_i|}}{|\mathbf{r}-\mathbf{r}_i|}, \tag{9}$$

where

$$p_{\mathrm{inc}}^i = \left(p_{\mathrm{inc}}^0(\mathbf{r}_i) + \sum_{j\neq i, j=1}^{N} A_j\,\frac{e^{ik|\mathbf{r}_i-\mathbf{r}_j|}}{|\mathbf{r}_i-\mathbf{r}_j|}\right)$$

is the total wave field incident on the scatterer. This total field is the sum of the incident field when there are no other scatterers and the scattered waves from the other scatterers. Note that $\hat{f}_i\, p_{\mathrm{inc}}^i$ is only a symbolic notation to indicate the response of the scatterer to the incident wave. Equating Eqs. (8) and (9) we obtain

$$A_i = \hat{f}_i\left(p_{\mathrm{inc}}^0(\mathbf{r}_i) + \sum_{j\neq i, j=1}^{N} A_j\,\frac{e^{ik|\mathbf{r}_i-\mathbf{r}_j|}}{|\mathbf{r}_i-\mathbf{r}_j|}\right). \tag{10}$$

Applying the same procedure to the other $N-1$ scatterers yields a system of N equations for the N unknown coefficients A_i. These equations were originally derived by Foldy,[13] and later summarized by Ishimaru.[8] By solving them simultaneously, the values of all the coefficients may be determined.

Now let us return to the problem at hand. The total wave field incident on the air bubble consists of three components: (a) the direct wave from the acoustic source (p_i^0); (b) the reflection from the surface of the direct wave from the source (p_r); and (c) the wave scattered back and forth between the bubble and the surface (p_m), which we will represent here by a multiple-scattering process between the bubble and its image. The first two waves can be represented as[12]

$$p_i^0 = e^{ik\hat{k}_i\cdot\mathbf{r}_1}, \tag{11}$$

and[14]

$$p_r = -e^{ik\hat{k}_r\cdot\mathbf{r}_1}, \tag{12}$$

where $\hat{k}_r$ is the unit vector for the direction of reflection, whose Cartesian components are given by

$$(\hat{k}_r)_x = (\hat{k}_i)_x, \quad (\hat{k}_r)_y = (\hat{k}_i)_y, \quad (\hat{k}_r)_z = -(\hat{k}_i)_z .$$

From Eq. (7), the incident wave due to the multiple scattering is clearly

$$p_m = -A\,\frac{e^{ik|\mathbf{r}_1-\mathbf{r}_2|}}{|\mathbf{r}_1-\mathbf{r}_2|} = -A\,\frac{e^{2ikd}}{2d}. \tag{13}$$

The total field incident on the bubble is then given by a sum of contributions from Eqs. (11), (12), and (13) as

$$p_i = e^{ik\hat{k}_i\cdot\mathbf{r}_1} - e^{ik\hat{k}_r\cdot\mathbf{r}_1} - A\,\frac{e^{2ikd}}{2d}, \tag{14}$$

and this aggregate field induces scattering from the air bubble as prescribed by Eq. (1), i.e.,

$$p_s(\mathbf{r}) = f\left(e^{ik\hat{k}_i\cdot\mathbf{r}_1} - e^{ik\hat{k}_r\cdot\mathbf{r}_1} - A\,\frac{e^{2ikd}}{2d}\right)\frac{e^{ik|\mathbf{r}-\mathbf{r}_1|}}{|\mathbf{r}-\mathbf{r}_1|}. \tag{15}$$

Self-consistency requires that Eq. (15) be equated to Eq. (6), so that

$$f\left(e^{ik\hat{k}_i\cdot\mathbf{r}_1} - e^{ik\hat{k}_r\cdot\mathbf{r}_1} - A\,\frac{e^{2ikd}}{2d}\right)\frac{e^{ik|\mathbf{r}-\mathbf{r}_1|}}{|\mathbf{r}-\mathbf{r}_1|} = A\,\frac{e^{ik|\mathbf{r}-\mathbf{r}_1|}}{|\mathbf{r}-\mathbf{r}_1|}, \tag{16}$$

allowing the unknown coefficient A to be readily determined, i.e.,

$$A = \frac{f(e^{ik\hat{k}_i\cdot\mathbf{r}_1} - e^{ik\hat{k}_r\cdot\mathbf{r}_1})}{1 + f(e^{2ikd}/2d)}. \tag{17}$$

Substituting from Eq. (3), we find

$$A = \frac{a(e^{ik\hat{k}_i\cdot\mathbf{r}_1} - e^{ik\hat{k}_r\cdot\mathbf{r}_1})}{\omega_0^2/\omega^2 - 1 - i\,\delta + (a/2d)e^{2ikd}}, \tag{18}$$

which is similar to the results found previously by Tolstoy[3] and Twersky.[7]

Feuillade[9] has developed an alternative self-consistent approach to solve this problem. In his method, the scattering of sound by a spherical air bubble near a boundary is described by using the concept of coupled oscillators. The bubble is again driven into oscillation by three field components which are, as before: (a) the direct wave (p_i^0); (b) the reflected wave (p_r); and (c) the wave scattered from the bubble and reflected back to itself by the surface (p_m). In his approach, however, the effect of the total field is described by incorporating these three components additively as a composite driving term in Devin's ''mass-spring'' equation for the monopole resonance of a bubble.[15] Including the time-variance $e^{-i\omega t}$ in the field terms, this procedure leads to the following expression:

$$\frac{\rho_l}{4\pi a}\{\ddot{v}+\omega\delta\dot{v}+\omega_2^0 v\}=-\left[(e^{ik\hat{k}_i\cdot\mathbf{r}_1}-e^{ik\hat{k}_r\cdot\mathbf{r}_1})e^{-i\omega t}\pm\frac{\rho_l e^{2ikd}}{8\pi d}\ddot{v}\right]. \tag{19}$$

On the left-hand side (LHS) of this equation, v is the differential volume (the difference between the instantaneous and equilibrium bubble volumes). This parameter is defined to represent the response of the bubble to the total incident pressure field [i.e., all the terms on the right-hand side (RHS)]. The RHS terms are all preceded by a minus sign, since a decrease in pressure results in an increase in the bubble volume. The first two terms are just p_i^0 and p_r, as defined by Eqs. (11) and (12), respectively, with time-variance included. The third term represents the influence on the bubble of its own scattered field after reflection in the surface. Note that this term contains $\ddot{v}$,[9,16] which also appears on the LHS, and that its effect is modified by a phase factor (i.e., $\pm$), which depends on whether the surface is rigid (''+'') or free (''−''). After choosing the ''−'' sign for this particular application, the next stage follows upon recognizing that both terms involving $\ddot{v}$ can be incorporated on the LHS. This yields a new expression, i.e.,

$$\frac{\rho_l}{4\pi a}\left\{\left(1-\frac{ae^{2ikd}}{2d}\right)\ddot{v}+\omega\delta\dot{v}+\omega_0^2 v\right\}=-[e^{ik\hat{k}_i\cdot\mathbf{r}_1}-e^{ik\hat{k}_r\cdot\mathbf{r}_1}]e^{-i\omega t}. \tag{20}$$

The advantage of formulating the problem this way becomes immediately clear. The interaction of the bubble with its own scattered field leads to a simple 2nd-order equation, with a modified bubble ''mass'' term on the LHS [cf. Eq. (19)], and only two driving terms on the RHS due to the direct wave (p_i^0) and the reflected wave (p_r). Assuming a harmonic solution $v=\overline{v}e^{-i\omega t}$, and substituting this in Eq. (20), gives the following expression for the time-independent differential volume $\overline{v}$:

$$\overline{v}=\frac{-(4\pi a/\omega^2\rho_l)(e^{ik\hat{k}_i\cdot\mathbf{r}_1}-e^{ik\hat{k}_r\cdot\mathbf{r}_1})}{\omega_0^2/\omega^2-1-i\delta+(a/2d)e^{2ikd}}. \tag{21}$$

In this formulation $A=(\rho_l/4\pi)\ddot{v}=-\omega^2(\rho_l/4\pi)\overline{v}$, and substitution for $\overline{v}$ from Eq. (21), gives back A identically as in Eq. (18).

From Eq. (18), we see that as $d\to\infty$, A becomes

$$A=\frac{a(e^{ik\hat{k}_i\cdot\mathbf{r}_1}-e^{ik\hat{k}_r\cdot\mathbf{r}_1})}{\omega_0^2/\omega^2-1-i\delta},$$

and taking this into Eq. (14), we find that the incident wave on the bubble is $e^{ik\hat{k}_i\cdot\mathbf{r}_1}-e^{ik\hat{k}_r\cdot\mathbf{r}_1}$, i.e., the simple sum of the direct wave and the reflected wave, indicating that the interaction between the bubble and the surface is negligible. This is the expected result as d goes to infinity. It is also intuitively clear from Eq. (20), since the modification to the ''mass'' term on the LHS also disappears when d becomes very large, and the equation reduces to that of an uncoupled bubble in free space, driven by the direct and reflected waves as represented by the two terms on the RHS.

When the air bubble is near the surface, the total scattered field is the sum of the waves from the air bubble and its image, i.e.,

$$p_{\text{scat}}(\mathbf{r})=A\left(\frac{e^{ik|\mathbf{r}-\mathbf{r}_1|}}{|\mathbf{r}-\mathbf{r}_1|}-\frac{e^{ik|\mathbf{r}-\mathbf{r}_2|}}{|\mathbf{r}-\mathbf{r}_2|}\right). \tag{22}$$

Then, substituting the general expression for A from (18) into this equation, we obtain

$$p_{\text{scat}}(\mathbf{r})=\frac{a(e^{ik\hat{k}_i\cdot\mathbf{r}_1}-e^{ik\hat{k}_r\cdot\mathbf{r}_1})}{\omega_0^2/\omega^2-1-i\delta+(a/2d)e^{2ikd}}\times\left(\frac{e^{ik|\mathbf{r}-\mathbf{r}_1|}}{|\mathbf{r}-\mathbf{r}_1|}-\frac{e^{ik|\mathbf{r}-\mathbf{r}_2|}}{|\mathbf{r}-\mathbf{r}_2|}\right). \tag{23}$$

This represents the complete solution for the scattered wave from a bubble near a free surface.

In the far field, where $r\to\infty$ (so that $|\mathbf{r}-\mathbf{r}_{1,2}|\approx r-\mathbf{r}\cdot\mathbf{r}_{1,2}/r$), the total scattered wave can be written in a Sommerfeld-radiation form, i.e.,

$$p_{\text{scat}}(\mathbf{r})=A(e^{-ik\hat{s}\cdot\mathbf{r}_1}-e^{-ik\hat{s}\cdot\mathbf{r}_2})\frac{e^{ikr}}{r}, \tag{24}$$

where $\hat{s}$ denotes the unit vector for the scattering direction, and A is again given by Eq. (18). To explore the features of the scattered wave, it is convenient to define an effective scattering amplitude $[F(\hat{k}_i,\hat{k}_s)]$ for the air bubble near the boundary as

$$p_{\text{scat}}(\mathbf{r})\equiv F(\hat{k}_i,\hat{k}_s)e^{ik\hat{k}_i\cdot\mathbf{r}_1}\frac{e^{ik|\mathbf{r}-\mathbf{r}_1|}}{|\mathbf{r}-\mathbf{r}_1|}\approx F(\hat{k}_i,\hat{k}_s)e^{ik\hat{k}_i\cdot\mathbf{r}_1-ik\hat{k}_s\cdot\mathbf{r}_1}\frac{e^{ikr}}{r}, \tag{25}$$

which is a definition consistent with Eq. (29) of Ref. 4.

Equating Eqs. (24) and (25) yields the effective scattering amplitude, i.e.,

$$F(\hat{k}_i,\hat{k}_s)=A\,\frac{(e^{-ik\hat{k}_s\cdot\mathbf{r}_1}-e^{-ik\hat{k}_s\cdot\mathbf{r}_2})}{e^{ik(\hat{k}_i-\hat{k}_s)\cdot\mathbf{r}_1}}$$

$$=\frac{a}{\omega_0^2/\omega^2-1-i\delta+(a/2d)e^{2ikd}}$$

$$\times(1-e^{ik\hat{k}_s\cdot(\mathbf{r}_1-\mathbf{r}_2)})(1-e^{ik(\hat{k}_r-\hat{k}_i)\cdot\mathbf{r}_1}). \quad (26)$$

This relation, which is valid for continuous plane-wave incidence, shows explicitly two effects of the pressure-release boundary: (a) the resonance frequency and damping of the single bubble are modified by the factor $(a/2d)e^{2ikd}$ [see Eqs. (14) and (15), respectively, of Ref. 9]; and (b) the scattering amplitude is modulated by the reflection from the surface. (N.B., the above equation does apply in the case of short beamed pulse incidences, and modifications must be made.)

The new resonance frequency (ω_R) is determined by equating the real part of the denominator in Eq. (26) to zero; and is the root of the equation

$$\omega_R^2=\frac{1}{1-(a/2d)\cos(2\omega_R d/c)}\,\omega_0^2. \quad (27)$$

When $2kd\ll 1$, this equation gives, to first order:

$$\omega_R^2=\frac{1}{1-a/2d}\,\omega_0^2, \quad (28)$$

which is in agreement with the approximate formula of Strasberg.[1] As the bubble moves away from the surface (i.e., as d increases), the resonance approaches that of a bubble in a full space.

Correspondingly, the new damping factor (δ_R) is given by

$$\delta_R=\delta-\frac{a}{2d}\sin(2kd). \quad (29)$$

Note the negative sign in front of the second term. In conjunction with Eq. (5), this expression shows that the effect of the boundary on the damping is to modify the radiation damping. To demonstrate this more explicitly, consider the case $2kd\ll 1$. Using the Taylor expansion, $\sin x\approx x-1/6x^3$, it is readily shown that, to first order in $2kd$, the original radiation damping factor (i.e., $\delta_r=ka$) is exactly canceled by the effect of the free surface. To third order in kd, the new radiation damping factor is $[2(kd)^2/3](ka)$. This is always smaller than ka and, for $kd\ll 1$, results in a significant reduction in radiation damping. This accords with the conjecture put forward by Strasberg[5] when trying to explain the results of Gaurnaurd and Huang (see the Introduction). It is possible that the presence of other sources of damping (i.e., δ_t and δ_v) provides the mechanism for limiting the magnitude of bubble pulsations at resonance. We shall examine this question in greater detail later.

Compared to the modal series solution of Gaunaurd and Huang,[4] the results in Eqs. (24) and (26) are much simpler, and demonstrate the physical significance of the boundary more clearly. In addition, the present results are easily checked, and may be readily compared with the results from other solutions involving complicated numerical computation. The modal series solution explicitly takes into account contributions from high-order modes, which could be important in certain cases. The self-consistent method discussed here has been specifically applied to a low-frequency case. It can be extended to higher frequencies, but in these cases the scattering function for a single bubble cannot be generally approximated as isotropic. The reader may refer to Twersky[7] for details. We will next examine the present results by several numerical examples.

II. THE NUMERICAL EXAMPLES

Let us first consider the effect of the surface on the resonance frequency. The frequency increase ratio, defined as ω_R/ω_0, can be determined from Eq. (27), and is the root of the following equation

$$x^2=\frac{1}{1-(a/2d)\cos(0.0272xd/a)}, \quad (30)$$

where we have used the well-known relation $\omega_0 a/c\approx 0.0136$ for air bubbles in water at a pressure of one atmosphere.[11] Note that the frequency increase x does not depend on the damping of the bubble, but only upon the value of $\omega_0 a/c$ and the distance from the surface. In Fig. 2(a), x is plotted against the ratio of the surface distance to radius (d/a), and generally compares well with the results of Strasberg.[1] However, at the special case $d=a$, i.e., when the bubble is immediately adjacent to the surface, the present theory predicts a frequency increase of 1.4137, which is smaller than the value of 1.5119 predicted by the approximate formula of Ref. 1. A more detailed comparison is presented in Table I, where the corresponding results given in Ref. 4 are also listed. Careful inspection of this table indicates very favorable agreement between the three methods for calculating the resonance frequency for surface distances $d=2a$ and greater. The largest difference (0.37%) occurs between the methods of Refs. 1 and 4 when $d=8a$. What is especially noticeable is close agreement (0.026% difference) between the values given by Eq. (27) and Ref. 4 when $d=2a$. This gives strong grounds for confidence that the monopole approximation adopted in the present paper is a good assumption for surface distances of this order and greater, and that it provides a sound basis for a comparison of the two methods. Table I shows close agreement between the three theoretical methods for values $d=2a$ and greater.

Now let us consider the scattering from the bubble. Although the theory presented here permits the study of scattering for arbitrary incident and scattering directions, we will concentrate on backscattering in this paper, and for this case $\hat{k}_s=-\hat{k}_i$. We will assume that the incident plane wave lies in the x-z plane, so that the incident and reflected unit vectors are

$$\hat{k}_i=(\sin\theta,0,\cos\theta),\quad \hat{k}_r=(\sin\theta,0,-\cos\theta),$$

where θ is the angle between the incident direction and the upward z axis. The scattering unit vector is then

$$\hat{k}_s=(-\sin\theta,0,-\cos\theta).$$

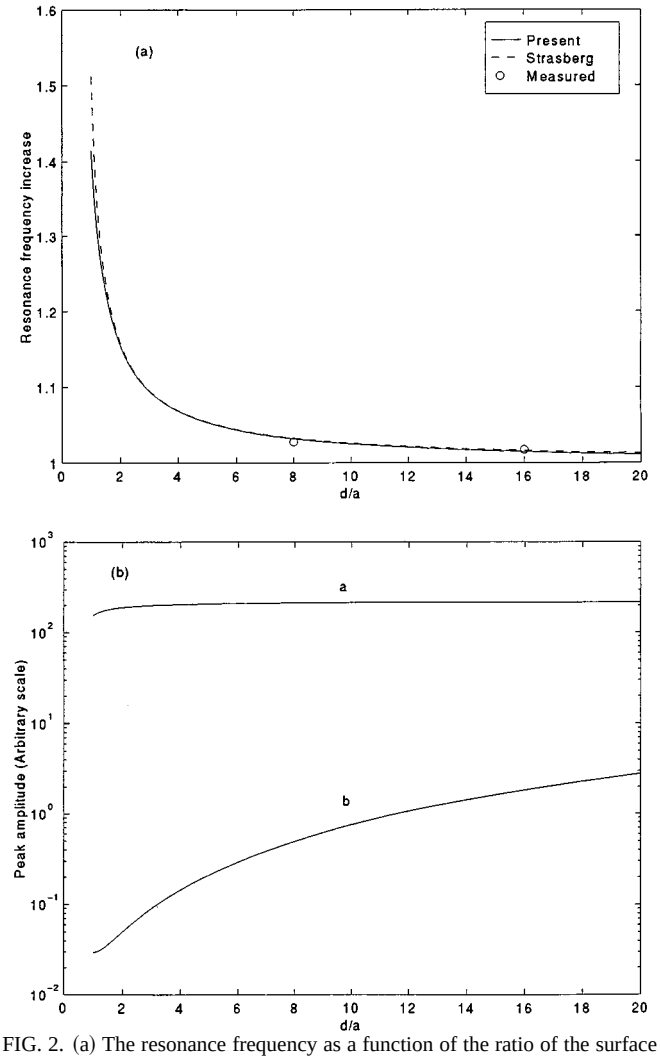


FIG. 2. (a) The resonance frequency as a function of the ratio of the surface distance to the bubble radius d/a; (b) The scattering amplitude at resonance as a function of d/a.

In the coordinate system of Fig. 1, the positions of the bubble and its image are

$$\mathbf{r}_1=(0,0,-d), \quad \mathbf{r}_2=(0,0,d).$$

Under these conditions, the effective scattering amplitude from Eq. (26) can be simplified as

$$F(\hat{k}_i,-\hat{k}_i)=\frac{a}{\omega_0^2/\omega^2-1-i\,\delta+(a/2d)e^{2ikd}} \times(1-e^{2ikd\,\cos\,\theta})^2. \tag{31}$$

TABLE I. Effect of the free surface on resonance frequency of a spherical air bubble.

d/a	Calculated[a]	Calculated[b]	Calculated[c]	Measured[d]
1	1.4137	1.5119		
2	1.1543	1.1577	1.154	
4	1.0685	1.0692	1.066	
8	1.0319	1.0328	1.029	1.028
16	1.0144	1.0160	1.015	1.017

[a]Results from Eq. (27).
[b]Theoretical results from Ref. 1.
[c]Calculated results from Ref. 4.
[d]The measured results cited in Ref. 1.

For convenience, we will define the reduced backscattering target strength as

$$\mathrm{TS}(\theta)\equiv 20\,\log(|F(\hat{k}_i,-\hat{k}_i)|/a), \tag{32}$$

where $|F(\hat{k}_i,-\hat{k}_i)|/a$ is given by

$$\frac{|F(\hat{k}_i,-\hat{k}_i)|}{a} =\frac{4\,\sin^2(kd\,\cos\,\theta)}{\sqrt{(\omega_0^2/\omega^2+(a/2d)\cos(2kd)-1)^2+(\delta-(a/2d)\sin(2kd))^2}}. \tag{33}$$

The oscillatory factor $\sin^2(kd\cos\theta)$ which modulates the scattering amplitude in the numerator of this expression is due to the conventional Lloyd's mirror interference effect.[17] The first null occurs at frequency $c/2d\,\cos\theta$, and the separation of the nulls is also $c/2d\,\cos\theta$.

At resonance, the scattering amplitude is

$$\frac{|F(\hat{k}_i,-\hat{k}_i)|}{a}=\frac{4\,\sin^2(kd\,\cos\,\theta)}{\delta-(a/2d)\sin(2kd)}. \tag{34}$$

When both the thermal and viscosity dampings are not included, which is the case considered in Ref. 4, this equation reduces to

$$\frac{|F(\hat{k}_i,-\hat{k}_i)|}{a}=\frac{4\,\sin^2(kd\,\cos\,\theta)}{ka-(a/2d)\sin(2kd)}. \tag{35}$$

Figure 2(b) shows the normalized $|F(\hat{k}_i,-\hat{k}_i)|/a$ at resonance as a function of d/a when $\theta=45°$. Curve a includes only radiation damping, while curve b incorporates radiation, thermal, and viscous damping effects. These plots show that the scattering amplitude at resonance decreases as the bubble moves towards the surface, as one would intuitively expect, since a nearby free surface converts the bubble monopole into an acoustic dipole with opposite strength, which radiates less efficiently. From these curves, it is clear that the scattering amplitude is reduced considerably by the inclusion of thermal and viscous damping. Consider the special case where thermal and viscous damping effects are not included. The result (represented by curve a) shows that the peak value decreases slightly as the bubble moves closer to the surface. This appears to conflict with the result obtained by Gaunaurd and Huang.[4] In his commentary on the paper of Gaunaurd and Huang, Strasberg[5] tried to explain their result by conjecturing that, in the absence of other sources of damping, the reduction in radiation damping due to the free surface might lead to a sufficiently large increase in the bubble pulsation amplitude at resonance to overcompensate for the reduced radiation efficiency. He suggested that this might result in a net increase in the calculated values of the scattering function. According to the present results the radiation damping is indeed reduced by the effect of the surface, but this reduction is not large enough to compensate for the reduced radiation efficiency. In fact, as the bubble moves closer to the surface, the driving pressure (i.e., the sum of the direct wave and its reflection from the surface) decreases as kd. The radiation from a dipole of opposite sign also

varies as kd. Together, these give rise to a radiation reduction which varies as $(kd)^2$. As discussed earlier, the reduced radiation damping is $\sim(kd)^2(ka)$ at resonance. Combining these two effects, we see that the peak value is proportional to $(ka)^{-1}$, i.e., it is proportional to the inverse of the resonance frequency. Now, as the bubble moves towards the surface, the resonance frequency will increase slightly so that, correspondingly, the peak value will decrease slightly. Why does this result differ from that of Gaunaurd and Huang? A quick response might be that, when the bubble is very close to the surface, the assumption of monopole scattering is invalid. There are, however, three reasons why this seems unlikely. First, the resonance feature itself is a direct result of the bubble's radially pulsating vibration. Second, while Fig. 2(b) plots $|F(\hat{k}_i,-\hat{k}_i)|/a$ in to a minimum surface distance $d=a$ (at which point, since the bubble touches the surface, the monopole approximation may indeed be questionable), the observed decrease in scattering amplitude actually begins at much greater values of d, and for $d=2a$ and greater Fig. 2(a) and Table I strongly indicate that monopole effects dominate the physical behavior of the system. Third, the results of Gaunaurd and Huang predict an enhancement of the scattering amplitude beginning at quite large values of the surface distance (e.g., $d=16a$), where monopole scattering is surely dominant. In addition, as we shall see below, the results given by the present method agree with those of the modal series solution at off-resonance frequencies, when the bubble is not too far away from the surface, and when thermal and viscous damping effects are not included.

In Fig. 3(a) the reduced target strength is plotted as a function of log(ka) for different values of d/a. The solid lines indicate results where radiation, thermal, and viscous damping effects are included. The dashed lines indicate results where only radiation damping is included. The radius of the bubble is set to 100 μm, while $c=1486$ m/s. The dotted line in Fig. 3(a) is for a bubble with no boundary present, and incorporates radiation, thermal, and viscous dampings. It is clear from the curves that the resonance frequency steadily increases as the bubble is moved towards the surface, whereas the peak strength at resonance shows a corresponding progressive decrease. We also see that the peak amplitude is reduced considerably by the inclusion of thermal and viscous damping effects, and that the peak scattering strength of a bubble is then smaller than the corresponding case where no boundary is present. These results correct an error in Fig. 10 of Ref. 9, which arose from a sign error in Eq. (21) of that paper. The present results agree with those shown in Fig. 3(a) of Gaunaurd and Huang[4] for off-resonance frequencies, when thermal and viscous dampings are not included. The resonance frequencies from the two solutions are also in agreement.

In Fig. 3(b), the reduced target strength is plotted over a wider range of log(ka) for the case $d/a=4$. The dashed lines again show the results obtained when thermal and viscous damping effects are not included. This figure indicates that the scattering strength steadily increases as the frequency is increased beyond resonance, that it becomes greater than the scattering strength of a bubble with no boundary present, and then drops after reaching a maximum. We can clearly see

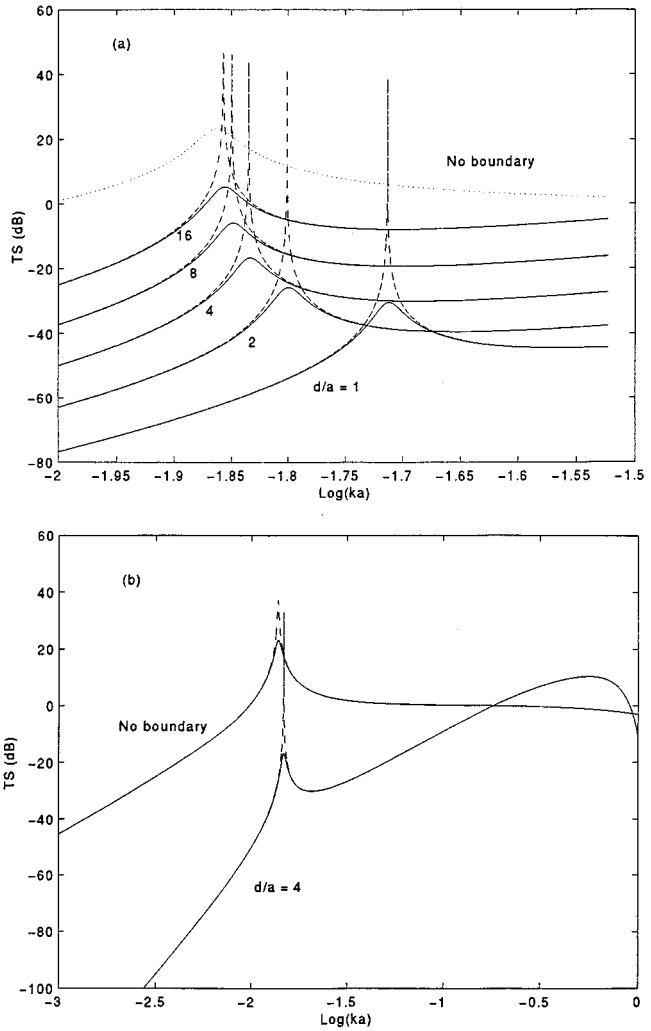


FIG. 3. The reduced backscattering target strength from Eq. (32) as a function of log(ka): (a) The target strength is plotted for several values of d/a. The dashed lines indicate results where thermal and viscous damping effects are not included; the dotted line indicates the case where no surface is present. $\theta=45°$. (b) The target strength is plotted over a broader range of ka for $d/a=4$. Again, the dashed lines indicate results where thermal and viscous damping effects are not included. $\theta=45°$.

that, without thermal and viscous damping, the scattering amplitude has a very sharp peak at the resonance.

In Fig. 4, the reduced target strength is plotted against ka for $d/a=50$, 100 and 500. Here we see that, as the bubble is moved further from the surface, regular oscillatory features appear in the scattering strength. The observed peaks and nulls are due to interference effects between the bubble and surface (or, equivalently, the bubble and its image). The nulls, which appear at regular frequency intervals, become more numerous and are spaced more closely together as the bubble is moved away from the surface. Each curve also contains a low-frequency peak (i.e., at $ka=0.0136$) due to the bubble resonance, and by noting the superposition of these peaks for the different cases we can see, when the bubble is moved far away from the surface, that the resonance frequency becomes essentially constant.

The frequency separation between the nulls in Fig. 4 is determined solely by the value of $kd\cos\theta$, as predicted by the theory. When the bubble is far away from the surface, the

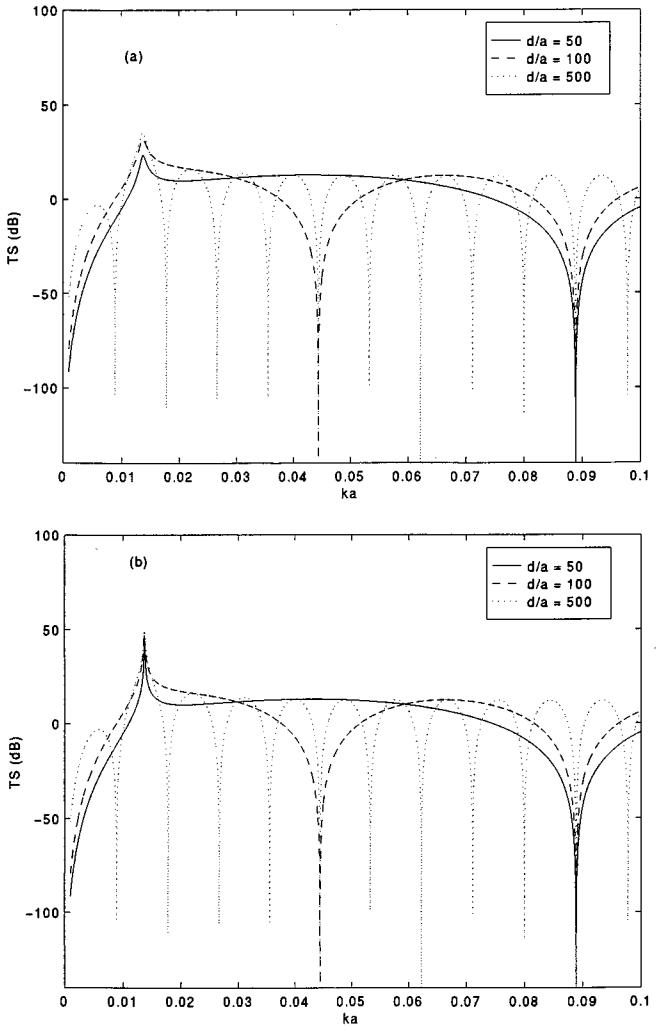


FIG. 4. The reduced backscattering target strength from Eq. (32) as a function of ka for larger values of d/a: (a) radiation, thermal, and viscous damping effects are included; (b) thermal and viscous damping effects are not included. $\theta = 45°$.

multiple-scattering interaction between the bubble and the surface is negligible, and the total driving pressure on the bubble is simply the sum of the direct incident wave $e^{ik\hat{k}_i \cdot \mathbf{r}_1}$ and the reflected wave from the surface $-e^{ik\hat{k}_r \cdot \mathbf{r}_1}$. The phase difference between these two waves is $2kd \cos\theta$, and another π phase shift must be added due to reflection in the pressure-release surface. Likewise, the total field scattered from the bubble is given by coherently adding the wave scattered directly from the bubble, and that reflected by the surface. This introduces another phase difference $2kd \cos\theta$, plus a further π phase due to the surface. Overall, a phase factor $(1-e^{2ikd\cos\theta})^2$ is introduced, as denoted in Eq. (31). This factor gives rise to the regular interference pattern in Fig. 4, which is due to the Lloyd's mirror effect.[17] By inspecting the various curves in Fig. 4 it is clear that, at the middle and higher frequencies, the target strength for a bubble reflected in a surface oscillates above and below the asymptotic value for a bubble when no surface is present, as shown in Fig. 3(a). Comparison of Fig. 4 with Fig. 3(b) of Gaunaurd and Huang[4] shows that, while they also observed oscillations in scattering strength, the spacing of their nulls was not governed by the $kd \cos\theta$ phase factor evident in the present work, and therefore did not exhibit the expected Lloyd's mirror behavior. Their curves also indicate significantly fewer oscillations and, in particular, an interference null, which appears below the resonance frequency for the $d/a = 500$ case in the present work, is completely absent from their figure.

III. CONCLUSIONS

We have considered acoustic scattering by a spherical air bubble near a flat pressure-release surface. Using the method of images, and an elegant self-consistent procedure, a simple formula for the sound scattering function is derived in which the resonance frequency and damping factor are explicitly shown. Application of the formula indicates that, when the bubble is moved closer to the surface, the resonance frequency increases, while the scattering amplitude at resonance decreases. When the bubble is moved away from the boundary, regular oscillatory features appear as a result of the interference between the bubble and the surface. The results from this study are compared with previous theoretical results and experimental observations. Agreements and disagreements are illustrated through numerical examples.

ACKNOWLEDGMENTS

The authors would like to thank Dr. M. Strasberg and Dr. D. Farmer for helpful discussions. This work was supported by DFO, Canada, and the US Office of Naval Research.

[1] M. Strasberg, ''The pulsation frequency of non-spherical gas bubbles in liquids,'' J. Acoust. Soc. Am. **25**, 536–537 (1953).
[2] H. N. Oguz and A. Prosperetti, ''Bubble oscillation in the vicinity of a nearly plane free surface,'' J. Acoust. Soc. Am. **87**, 2085–2092 (1990).
[3] I. Tolstoy, ''Superresonant systems of scatterers. I,'' J. Acoust. Soc. Am. **80**, 282–294 (1986).
[4] G. C. Gaunaurd and H. Huang, ''Acoustic scattering by an air-bubble near the sea surface,'' IEEE J. Ocean Eng. **20**, 285–292 (1995).
[5] M. Strasberg, ''Comments on 'Acoustic scattering by an air bubble near the sea surface','' IEEE J. Ocean Eng. **21**, 233 (1996).
[6] G. C. Gaunaurd and H. Huang, ''Reply to ''Comments on 'Acoustic scattering by an air bubble near the sea surface' '','' IEEE J. Ocean Eng. **21**, 233 (1996).
[7] V. Twersky, ''Multiple scattering of waves and optical phenomena,'' J. Opt. Soc. Am. **52**, 145–171 (1962).
[8] A. Ishimaru, *Wave Propagation and Scattering in Random Media* (Academic, New York, 1978), Chap. 14, Vol. 2.
[9] C. Feuillade, ''Scattering from collective modes of air bubbles in water and the physical mechanism of superresonances,'' J. Acoust. Soc. Am. **98**, 1178–1190 (1995).
[10] Z. Ye, ''Low-frequency acoustic scattering by gas-filled prolate spheroids in liquids,'' J. Acoust. Soc. Am. **101**, 1945–1952 (1997).
[11] C. S. Clay and H. Medwin, *Acoustical Oceanography* (Wiley, New York, 1977).
[12] L. M. Brekhovskih, *Waves in Layered Media* (Academic, New York, 1980).
[13] L. L. Foldy, ''The multiple scattering of waves. I. General theory of isotropic scattering by randomly distributed scatterers,'' Phys. Rev. **67**, 107–119 (1945).
[14] G. V. Frisk, *Ocean and Seabed Acoustics* (Prentice-Hall, Englewood Cliffs, NJ, 1994), pp. 36–37.
[15] C. Devin, ''Survey of thermal, radiation, and viscous damping of pulsating air bubbles in water,'' J. Acoust. Soc. Am. **31**, 1654–1667 (1959).
[16] H. Lamb, *Hydrodynamics* (Dover, New York, 1945), 6th ed., Sec. 56.
[17] R. J. Urick, *Principles of Underwater Sound* (McGraw-Hill, New York, 1983), 3rd ed., pp. 131–134.

High-frequency bistatic scattering by sub-bottom gas bubbles

Dezhang Chu
Department of Applied Ocean Physics and Engineering, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543

Kevin L. Williams, Dajun Tang, and Darrell R. Jackson
Applied Physics Laboratory, University of Washington, Seattle, Washington 98105

(Received 5 March 1996; accepted for publication 24 March 1997)

A previous study of high-frequency acoustic backscattering data collected at Eckernfoerde Bay, Germany revealed that scattering is mainly due to methane gas bubbles buried about a meter beneath the seafloor [Tang *et al.*, J. Acoust. Soc. Am. **96**, 2930–2936 (1994)]. A backscattering model was developed [Tang, Geo-Marine Lett. **16**, 161–169 (1996)] where the gas bubbles were approximated by oblate spheroids. In this paper, a bistatic scattering model is proposed as an extension of the previously developed backscattering model. In this model, gas bubbles are again assumed to be oblate spheroids with varying aspect ratios and a single-scattering approximation is used. The model is compared to bistatic data acquired in Eckernfoerde Bay, Germany. In particular, the azimuthal dependence of the bistatic scattering strength predicted by the model is tested against experimental data and it is found that both the model and the bistatic scattering strength data exhibit a mild azimuthal dependence. Best agreement between model and data requires a 35% reduction in areal bubble density relative to that used in the backscattering model/data comparison. Possible reasons for this are discussed including multiple scattering effects. [S0001-4966(97)04307-5]

PACS numbers: 43.30.Ft, 43.30.Gv, 43.30.Ma [JHM]

INTRODUCTION

Sub-bottom contributions to acoustic scattering have received increased attention in recent years. It has been shown in a number of studies that sub-bottom inhomogeneities can dominate acoustic backscattering at frequencies as high as 40 kHz.[1,2] However, the previous investigations have been restricted to monostatic scattering (backscattering) in which the dependence of the scattering directivity pattern on receiving grazing and azimuthal angles cannot be examined. Although such a scattering directivity pattern can lead to further insight into the scattering physics, very few bistatic experiments of the type required to examine the scattering directivity pattern have been reported due to the engineering and operational difficulties involved in data acquisition.

Likewise, development of bistatic models is a more demanding task. Recently, Jackson[3] developed a bistatic model including surface scattering from the rough water–sediment interface and volume scattering from the sub-bottom. The comparison of his model with experimental data shows a reasonable agreement.[4] However, this model treats volume inhomogeneities via perturbation theory and, as such, does not address in a fundamental way environments in which scattering is due to high contrast scatterers such as bubbles. Just such an environment was encountered in Eckernfoerde Bay, Germany during the Coastal Benthic Boundary Layer Special Research Program (CBBL-SRP). Within the Eckernfoerde mud sediment, a layer consisting of many nonspherical methane gas bubbles was found at about 1 m below the smooth seafloor.[5] These bubbles scatter sound far more efficiently than the water/seafloor interface and the mud above the bubble layer.[1,2] In this paper, a two-layer bistatic scattering model is proposed to describe scattering by such a bubble layer and compared with data taken at the Eckernfoerde site. Section I outlines the theoretical development; Section II presents some results of numerical simulations, compares theory with experimental data, and ends with some discussion of possible multiple scattering effects. Conclusions are drawn in Sec. III.

I. BISTATIC SCATTERING MODEL

Bistatic scattering experiments employ configurations in which the transmitter and the receiver are at different positions. The commonly used monostatic scattering or backscattering geometry is a special case of the more general bistatic scattering geometry. Similar to backscattering strength, bistatic scattering strength can be defined as the decibel equivalent of the scattering cross section per unit area per unit solid angle[3] and is, in general, a function of incident angles (θ_i, ϕ_i) and scattering angles (θ_s, ϕ_s), where θ and ϕ correspond to grazing and azimuth angles, respectively. When $\phi_s = \pi + \phi_i$, and $\theta_s = \theta_i$, bistatic scattering reduces to backscattering. Obviously, bistatic scattering measurements provide more information than backscattering measurements.

Previous backscattering studies at the Eckernfoerde Bay site showed that backscattering at 40 kHz is due to a layer of methane gas bubbles located about 1 m below the seafloor, and the scattering by these bubbles has been successfully modeled by a backscattering model where the gas bubbles were approximated by nonresonant oblate spheroids.[2] In addition to these backscattering data, bistatic scattering data were also collected with the same source and a mobile receiver at the same time and location.[4] The purpose of this

paper is to extend the backscattering model to the bistatic scenario.

Before introducing the model, we restate several assumptions made in the backscattering model that will be used also in the bistatic model. First of all, in Refs. 1 and 2, the surface scattering at 40 kHz from the seafloor itself was found insignificant and could be ignored due to the weak acoustic impedance change across the seafloor/water interface.

Second, it is observed that the bubbles are concentrated in a layer distributed over a vertical distance of about 10 cm. The differences between the grazing angles incident upon/scattered from the top and bottom of the layer are small (less than 2° for the geometry of the experiment). Since we consider single scattering only, it is reasonable to assume that all bubbles are on a surface with an equivalent bubble surface density obtained from integrating the volume density over depth. As a result, volume scattering is treated as quasi-surface scattering and the propagation loss due to volume scattering is ignored. In contrast, the horizontal distribution of the bubbles is assumed to be homogeneous, i.e., a uniform probability density function (PDF).

Third, x-ray tomographic analysis of the core data[6] shows that the shapes of the bubbles are nonspherical and their sizes vary from about 0.5 mm to less than 15 mm in volume-equivalent spherical radius. Smaller bubbles were not measured because of the limited resolution of the x-ray system. The bubbles are oriented as coins standing on their edges. Their broadside normals are, in general, parallel to the plane on which the bubbles are standing. For brevity, we call such a configuration *horizontally oriented*. To simplify our analysis and to objectively model bistatic scattering as in the backscattering model, we assume that resonant scattering is unimportant and can be neglected. In the present model, we assume that the bubbles can be properly modeled as horizontally oriented pressure-release oblate spheroids with various aspect ratios.

Finally, in the backscattering model, the effect of the multiple scattering among bubbles was not considered. Likewise, in the present bistatic scattering model multiple scattering will be ignored. This approximation will be revisited after model/data comparisons are presented.

A. Geometry of the two-layer model

A major issue of the two-layer model for a bistatic geometry is the determination of the ensonified region of the scattering layer. For a homogeneous medium, if the source signal is a short pulse, at any time greater than the minimum time required for the wave to travel from the source to the scattering layer and then to the receiver, the interception area on the scattering plane will be an elliptical ring whose width is proportional to the pulse length[4] and can be determined analytically.

However, for the two-layer model, since there are two media (water column and sediment) with different sound speeds, the acoustic wave changes its propagation direction when it passes through the interface that divides the two media. The interception area at a given time instant is no longer an elliptical ring and its shape cannot be determined

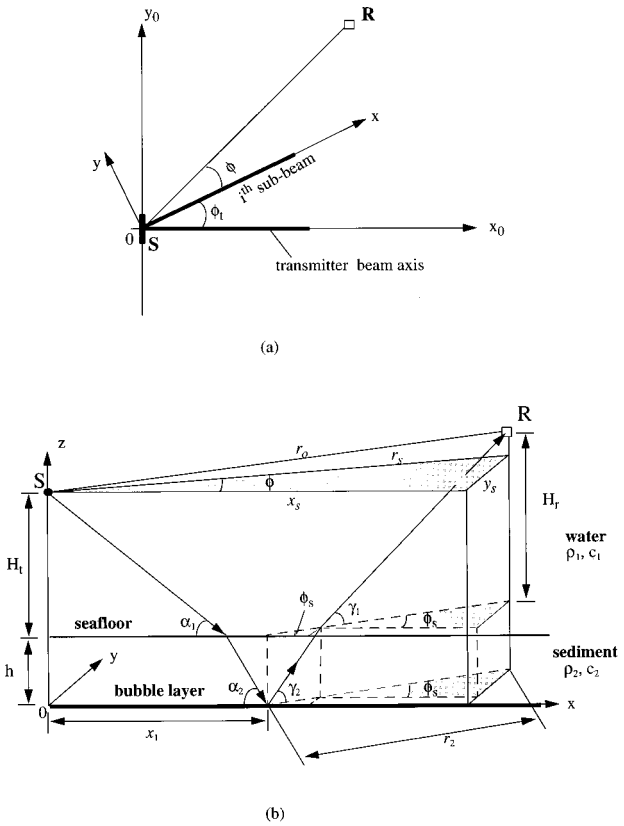


FIG. 1. Geometry of a bistatic scattering model. (a) Top view, where (x_0,y_0) are original coordinates and (x,y,z) are rotated coordinates. *S* and *R* stand for source and receiver, respectively. (b) Side view of the *i*th sub-beam.

analytically. To solve the problem numerically, we divide the transmit beam into a number of narrow, nonoverlapping sub-beams propagating in different directions, i.e., different azimuth angle ϕ_t, and compute the scattering contribution due to each individual sub-beam. Since multiple scattering is not considered in the current model, the total scattered field can be obtained by simply summing up the contributions from all these sub-beams weighted by the transmitter and receiver beampatterns.

To compute the scattering due to each individual transmit sub-beam, a global coordinate system is chosen in such a way that the the acoustic source or transmitter is located at $x=0$, the z axis is pointing up, and the bubble layer is at $z=0$ as shown in Fig. 1, where Fig. 1(a) is the top view and Fig. 1(b) the side view. In Fig. 1(a) there are two coordinate systems: the original and rotated coordinate systems. Their z axes coincide with each other pointing outward from the paper. The original coordinate system is chosen such that the x_0 axis coincides with the transmit beam axis, corresponding to zero azimuth angle. For a transmit sub-beam in an azimuthal direction ϕ_t (the angle between transmit beam axis and the sub-beam direction), the original coordinate system is rotated about the z axis by an angle of ϕ_t to make the transmit azimuth angle zero in the rotated coordinate system (x,y,z). The bistatic scattering geometry shown in Fig. 1(b) is the side view of the *i*th transmit sub-beam. The transmitter is located at $(0,0,H_t+h)$ and the chosen sub-beam is in the

x-z plane, while the receiver position can be out of the x-z plane and located at $(x_s, y_s, h+H_r)$, where H_t and H_r are distances from the seafloor to the transmitter and receiver correspondingly, and h is the thickness of the sediment overlying the scatterers. From Fig. 1 the following geometric relations can be easily obtained:

$$
\begin{aligned}
&r_s = \sqrt{r_0^2 - (H_r - H_t)^2}, \\
&x_s = r_s \cos\phi, \quad y_s = r_s \sin\phi, \\
&x_1 = H_t \cot\alpha_1 + h \cot\alpha_2, \\
&r_2 = H_r \cot\gamma_1 + h \cot\gamma_2, \\
&\phi_s = \sin^{-1}(y_s / r_2), \quad x_2 = r_2 \cos\phi_s,
\end{aligned}
\tag{1}
$$

where r_0 is the distance between the transmitter and the receiver, r_s the projection distance of r_0 to the x-y plane, α_1 and γ_1 are grazing angles at the seafloor, while α_2 and γ_2 are grazing angles at the bubble layer. Note that α_1 and α_2 correspond to the transmitted wave while γ_1 and γ_2 correspond to the scattered wave. Using the above geometrical relations, we have

$$
\begin{aligned}
x_s &= x_1 + x_2 \\
&= H_t \cot\alpha_1 + h \cot\alpha_2 + (H_r \cot\gamma_1 + h \cot\gamma_2) \\
&\quad \times \cos[\sin^{-1}(y_s / (H_r \cot\gamma_1 + h \cot\gamma_2))].
\end{aligned}
\tag{2}
$$

A constraint on the arrival time gives

$$
t_s = \frac{1}{c_1}\left(\frac{H_t}{\sin\alpha_1} + \frac{H_r}{\sin\gamma_1}\right) + \frac{h}{c_2}\left(\frac{1}{\sin\alpha_2} + \frac{1}{\sin\gamma_2}\right), \tag{3}
$$

where c_1 and c_2 are sound speeds of the water column and the sediment, respectively.

Grazing angles α_1, α_2, γ_1, and γ_2 in above equations are related by Snell's law:

$$
\cos\alpha_2 = \frac{c_2}{c_1}\cos\alpha_1, \tag{4}
$$

$$
\cos\gamma_2 = \frac{c_2}{c_1}\cos\gamma_1. \tag{5}
$$

For a given time, there are only two independent unknowns in Eqs. (2) and (3): γ_1 and α_1. These two simultaneous transcendental equations can be solved numerically. One way to solve the equations is to step through time intervals; N discrete time steps, for example. It requires solving the transcendental equations N times, which is not an efficient way.

An alternative way used in our modeling is to find all possible combinations of α_1 and γ_1 that satisfy Eq. (2) first, then compute t_s using Eq. (3) and the obtained α_1 and γ_1 from Eq. (2), and finally sort the result according to an ascending time sequence t_s. This way we only need to solve the transcendental equations once.

It can be proved that when the receiver is in the same plane as the transmit sub-beam $\phi = 0$, and when the scattering grazing angle γ_1 is equal to the transmit grazing angle α_1, the travel time reaches its minimum, $t_{\min}$, and this scattering grazing angle γ_1 can be considered as the grazing angle in the specular direction, $\gamma_{\rm sp}$. At any other time $t_s > t_{\min}$, there are always two rays intercepting the gas-bubble layer, i.e., the $z = 0$ plane: One corresponds to a grazing angle γ_1 larger than $\gamma_{\rm sp}$, and the other corresponds to a grazing angle smaller than $\gamma_{\rm sp}$. This bears some analogy to the case where only one homogeneous layer overlays the scattering layer and the ensonified area on the scattering layer is an elliptical ring. For any vertical plane containing the transmitting sub-beam and intercepting this scattering layer, there are always two intercepting points on the elliptical ring, corresponding to the two rays. For $\phi \neq 0$, there still exists a minimum travel time, $t_{\min}$, which corresponds to a single ray with the transmit and scattering grazing angles α_1 and γ_1. Here α_1 and γ_1 are not equal in general. At any time $t_s > t_{\min}$, there are still two rays with one ray having a scattering angle γ_1 greater than that corresponding to $t_s = t_{\min}$, and the other smaller than γ. The total scattering contribution at time t_s is the sum of these two rays.

B. Scattering model

As described in Refs. 1, 2, 5, and 6, the strong scattering layer consists of an aggregation of nonspherical gas bubbles which are modeled as (horizontally oriented) oblate spheroids with their major axis pointing randomly in directions orthogonal to the vertical. The model for bistatic scattering from these spheroids is an extension of the model developed by DiPerna and Stanton[7] in which a conformal mapping method is applied. Since the bubbles have different orientations and aspect ratios, it is reasonable to assume the oblate spheroidal scatterer has an omnidirectional orientation distribution in the x-y plane, i.e., bubbles are oriented uniformly over 0 to 2π. Obviously, for an oblate spheroid under such an assumption, the incident and azimuth angles with respect to the horizontal plane in the global coordinates or observation coordinates [rotated coordinates (x,y,z) in Fig. 1(b)] are different from those with respect to the spheroid defined in the local coordinates (shown in Fig. A1). To perform a statistical average over orientation, coordinate transformations are needed to relate the local coordinates to global coordinates. The details of the transformations are given in the Appendix A. The average cross section over orientation as well as aspect ratio expressed in the global coordinates can be obtained by

$$
\begin{aligned}
\mathscr{S}_s(\gamma_2, \phi_s; \alpha_2) = \frac{2}{\pi}\int_0^{\pi/2}\int_{-\infty}^{\infty} &\sigma_\epsilon(\Theta_s, \Phi_{si}; \Theta_i) \\
&\times \rho(\epsilon)\, d\phi_{\rm rot}\, d\epsilon,
\end{aligned}
\tag{6}
$$

where σ_ϵ is the scattering cross section of an oblate spheroid with an aspect ratio ϵ, Θ_s, Φ_{si}, and Θ_i are given by Eqs. (A11)–(A15) in Appendix A, and $\rho(\epsilon)$ is the probability density function (PDF) of the aspect ratio ϵ given in the next section.

Using $\mathscr{S}_s(\gamma_2, \phi_s; \alpha_2)$ together with α_1 and γ_1 obtained numerically using the method described in the previous subsection, the received acoustic intensity corresponding to the transmit azimuthal angle ϕ_t at time t for ray i can be expressed as

$$I_{si}(\gamma_2,\phi_s;\alpha_2,\phi_t)=\frac{I_0 e^{-\beta_w r_w} e^{-\beta_s r_{\rm sed}}}{r_t^2 r_{\rm scat}^2}\mathscr{S}_s(\gamma_2,\phi_s;\alpha_2)\times A\rho_b B_t(\theta_t,\phi_t)B_r(\theta_r,\phi_r)\times(T_{12}T_{21}), \tag{7}$$

where $i=1,2$ corresponding to the two rays $\gamma_1<\gamma_{\rm sp}$ and $\gamma_1>\gamma_{\rm sp}$, respectively,

$$r_t^2=\left(\frac{H_t}{\tan\alpha_1}+\frac{h}{\tan\alpha_2}\right)\left(\frac{H_t\tan\alpha_2}{\sin^2\alpha_1}+\frac{h\tan\alpha_1}{\sin^2\alpha_2}\right),$$
$$r_{\rm scat}^2=\left(\frac{H_r}{\tan\gamma_1}+\frac{h}{\tan\gamma_2}\right)\left(\frac{H_r\tan\gamma_2}{\sin^2\gamma_1}+\frac{h\tan\gamma_1}{\sin^2\gamma_2}\right) \tag{8}$$

are distances from the transmitter and the receiver to the scatterer, and

$$r_w=\frac{H_t}{\sin\alpha_1}+\frac{H_r}{\sin\gamma_1},$$
$$r_{\rm sed}=h\left(\frac{1}{\sin\alpha_2}+\frac{1}{\sin\gamma_2}\right) \tag{9}$$

are travel distances of a ray in the water and in the sediment, respectively. Here, β_w and β_s are attenuation coefficients in water and sediment, respectively (neper/m); T_{12} and T_{21} are energy transmission coefficients from medium 1 to 2 and 2 to 1; ρ_b is the bubble surface density (no./m^2); B_t and B_r are transmitter and receiver beampatterns with the arguments $\theta_t=\alpha_1-\theta_{\rm td}$, $\theta_r=\gamma_1-\theta_{\rm rd}$, and $\phi_r=\phi-\phi_{\rm rd}$, where $\theta_{\rm td}$ is the transmitter depression angle (mounting angle), and $\theta_{\rm rd}$ and $\phi_{\rm rd}$ are receiving depression and orientation angles accordingly. Here, A is the ensonification area and can be determined numerically (Appendix B). Note that γ_2, ϕ_s, α_2, and ϕ_t are functions of time. The total received intensity at time t, $I_s(t)$ can then be obtain by integrating I_{si} given by Eq. (7) over ϕ_t,

$$I_s(t)=\int_{-\pi/2}^{\pi/2}\sum_{i=1}^{2} I_{si}(\gamma_2,\phi_s;\alpha_2,\phi_t)d\phi_t, \tag{10}$$

where we have summed up the contributions from the two rays corresponding to the transmit direction ϕ_t. We ignored contributions for ϕ_t between $\pi/2$ and $3\pi/2$ due to the strong attenuation of the transmitter beampattern (greater than 40 dB). In Eq. (10), at a given time, the angles γ_2, ϕ_s, and α_2 can be uniquely determined if a transmit azimuthal angle ϕ_t is specified. Therefore, a summation over ϕ_t in Eq. (10) also results in a summation over γ_2, ϕ_s, and α_2.

In general, contributing rays with a variety of incident and scattered angles intercept the bubble layer at any given time; therefore one cannot obtain the conventional scattering strength defined by a pair of incident and scattered angles. However, by taking advantage of the narrow beamwidths of both transmitter and receiver, it is possible to determine whether the scattered field is dominated by the contribution from a single direction. The average scattering strength (SS) per unit area per unit solid angle can be expressed as

$$\mathrm{SS}(\gamma_2,\phi_s;\alpha_2)=10\log_{10}(\rho_b\mathscr{S}_s^{(d)}(\gamma_2,\phi_s;\alpha_2)), \tag{11}$$

where the superscript (d) denotes dominant scattering.

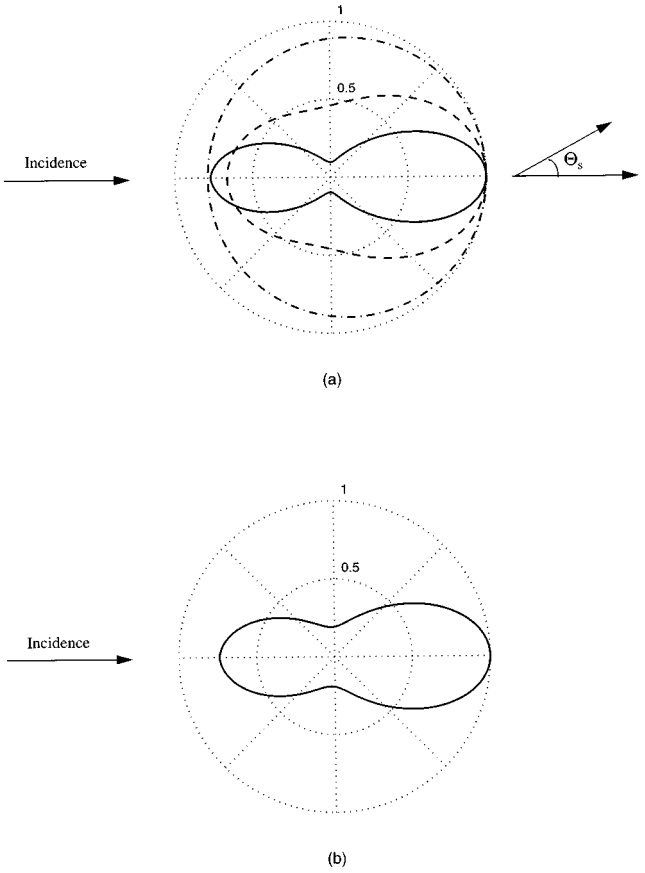


FIG. 2. Normalized bistatic scattering cross section of an oblate spheroid. In all computations, the frequency is 40 kHz, sound speed is 1425 m/s, and the semi-minor axis of the oblate spheroid is 1 mm, resulting in $ka=0.18$. (a) Scattering directivity pattern of an oblate spheroid with an aspect ratio of 1 (dash-dotted), 5 (dashed), and 10 (solid); (b) normalized average scattering cross section over aspect ratio using an exponential PDF. In both (a) and (b) the incidence is along the broadside normal of the oblate spheroid, i.e., $\Theta_i=0$ and the scattered angle Θ_s is the angle between the broadside normal and the scattered direction.

II. RESULTS AND DISCUSSIONS

A. Numerical simulations

At low frequencies, $ka\ll 1$, where k is the wave number and a is the characteristic geometric dimension of the object, scattering by a pressure-release sphere is almost omnidirectional. However it is directional for an oblate spheroid and approaches a dipolelike bistatic scattering pattern as aspect ratio increases. Figure 2(a) shows the normalized differential scattering cross section for a single oblate spheroid versus scattering angle Θ_s for a broadside incidence $\Theta_i=0$, where Θ_i and Θ_s are defined in a local coordinate system described in Appendix A. Θ_i represents the angle between the broadside normal and the incident direction, and Θ_s represents the angle between the broadside normal and scattering direction. Since the scattering is independent of azimuth angle, Φ_s, only the scattering in an arbitrary azimuthal plane is illustrated. In the figure, $\Theta_s=0$ corresponds to forward scattering while $\Theta_s=180$ corresponds to backscattering. The three curves in Fig. 2(a) correspond to oblate spheroids with aspect ratio 1 (dashed-dotted), 5 (dashed), and 10 (solid), respectively. For the curve with aspect ratio of unity, i.e., a

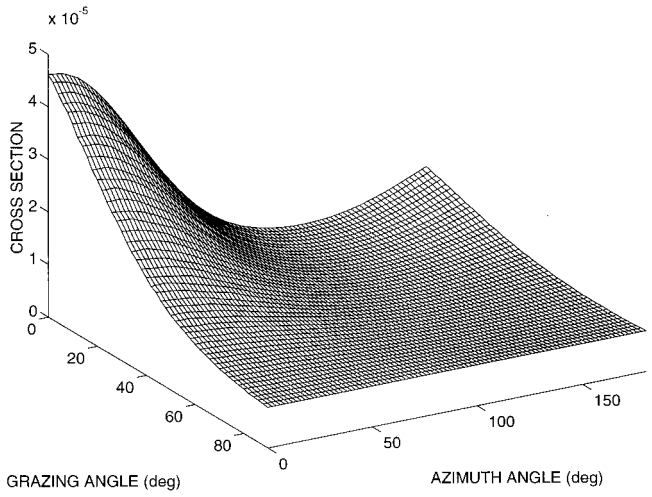


FIG. 3. Bistatic scattering cross section in the upper half-space for $\theta_i = 10°$.

sphere, the variation in scattering cross section is less than 20%, or 1 dB, while for the curve with aspect ratio of 10, the variation is greater than 80%, or 7 dB, and the cross section has a minimum at $\Theta_s = 90°$, showing a dipolelike scattering pattern. For the curve with aspect ratio 5, the variation is within the two extremes.

For an aggregation of oblate spheroidal bubbles with various aspect ratios, an average scattering cross section can be obtained using an appropriate PDF. To perform the average, we assume that the semi-minor axis of the oblate spheroid is kept unchanged, while the semi-major axis changes according to the PDF. From the core data,[8] it was observed that gas bubbles with smaller aspect ratios are more abundant than those with larger aspect ratios. A simple truncated exponential used in Ref. 2 is also used in our computation:

$$\rho(\epsilon) = \begin{cases} \dfrac{\beta}{e^{-\beta\epsilon_{\min}} - e^{-\beta\epsilon_{\max}}} e^{-\beta\epsilon}, & \epsilon_{\min} \leqslant \epsilon \leqslant \epsilon_{\max}, \\ 0, \quad \text{otherwise} \end{cases} \quad (12)$$

with $\beta = 0.25$, $\epsilon_{\min} = 1$, and $\epsilon_{\max} = 10$. Figure 2(b) illustrates the average scattering cross section normalized to unity and using such a truncated exponential PDF. It can be seen that, after averaging, the angular dependence has not been washed out entirely but is smoothed to some extent.

For a more general case, when the incidence is along an arbitrary direction (say θ_i as shown in Fig. A1) computations of the average bistatic cross section over orientations and aspect ratios given by Eq. (6) involve coordinate transformations described in Appendix A. Figure 3 shows a 3-D averaged bistatic scattering cross section in the upper half-space for $\theta_i = 18°$, where grazing and azimuth angles (θ_s, ϕ_s) vary from $0-90°$ and $0-180°$, respectively. Similar to the azimuth independent case shown in Fig. 2, the variation of the scattering cross section is about 8 dB. Again, it can be found that the maximum is in the forward direction while the minimum occurs at grazing angle $\theta_s = 90°$ when the azimuth angle ϕ_s is fixed or at azimuth angle $\phi_s = 90°$ when the grazing angle θ_s is fixed.

TABLE I. Parameters used the bistatic bottom scattering model.

$c_w = 1448$ m/s	sound speed in water
$c_s = 1425$ m/s	sound speed in sediment
$\rho_w = 1.0$ g/cm^3	water density
$\rho_s = 1.1$ g/cm^3	sediment density
$\alpha_w = 0$ dB/m	water attenuation coefficient
$\alpha_s = 2.4$ dB/m	sediment attenuation coefficient
$\rho_b = 8.8\times10^3$ (no./m^2)	equivalent bubble surface density

B. Comparison with bistatic experimental data

The 40-kHz bistatic scattering data were collected at Eckernfoerde Bay, Germany in April, 1993.[4] The Benthic Acoustic Measurement System (BAMS) served as the transmitter while the receiving array was suspended from a vessel that was moored at several locations in the vicinity of the source. The transmitter was mounted on a tripod 5 m above the seafloor, rotated 360° in 72 steps with a step size of 5°, and transmitted one pulse at each step. The receiver was about 7.5 m above the seafloor and was within a 50-m range of the transmitter for the data shown here. The transmitted signal was a 2-ms linear frequency modulated signal, with its frequency swept from 39 to 41 kHz. Beampatterns of both transmitter and receiver are narrower in the horizontal plane (about 6.5° beamwidth for both) and broader in the vertical plane (14° and 37° beamwidths for transmitter and receiver, respectively). Details of the system can be found in Ref. 4. Since the bistatic data were collected at the same location and the same time as the previously reported backscattering data,[1,2] the parameters used in backscattering model were initially kept unchanged in the current bistatic scattering model. However, in the final model/data comparison to be shown here the bubble concentration, ρ_b is reduced by about 35% (about 2 dB in scattering strength) to obtain the best fit. Possible reasons for this needed adjustment are discussed at the end of this section.

The parameters used in the modeling are listed in Table I, where all parameters are the same as those used in Refs. 1 and 2 except for bubble density discussed in the previous paragraph. The attenuation coefficient in the water column is set to zero since at 40 kHz for a range within 50 m, the attenuation *is* negligible.

In this study, a total of eight data sets have been analyzed. Each data set corresponds to one revolution (360°) of the transmitter (72 pings). Figure 4 shows the comparison of the reverberation level (RL) between the experimental data and the theoretical predictions for four pings from one data set, where the thinner lines are the experimental data and the thicker lines are the model predictions. These four pings are selected to represent the time series of the scattering from four different azimuthal directions and exhibit some distinct patterns. Since the transmission and reception systems used two independent internal clocks, the alignment of the data with the theoretical curves in time is adjusted primarily by aligning the first arrival of the data with the direct arrival predicted by the model based on the actual position and orientation information. For some pings, such adjustments could result in a misalignment between data and the theory, i.e., the major scattering structures (peaks and troughs) in the

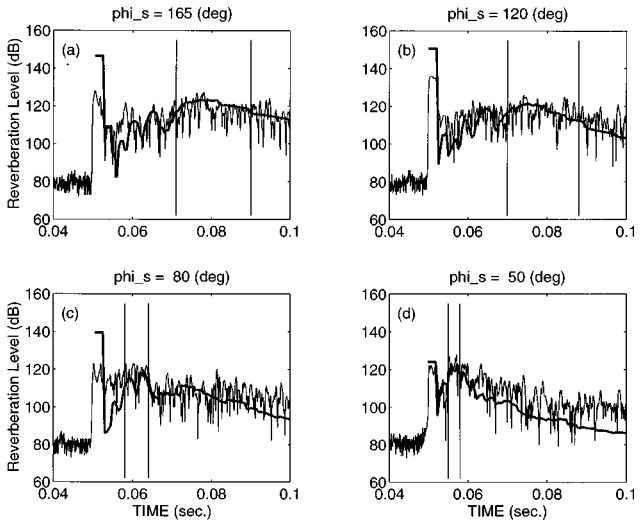


FIG. 4. Comparison of reverberation level between the bistatic scattering model predictions and the experimental data. The four time series are chosen from the same data set representing the bistatic scattering with four different azimuth angles as indicated in the figure. The vertical lines define the time windows within which the attenuation due to the beampattern (product of transmitting and receiving beampatterns) is less than 15 dB, and the data are considered ''reliable.''

RL time series can match much better if the RL curve is shifted a few ms. This could be attributed to the possible inaccuracy in determination of positions and orientations of the transducers.

From Fig. 4 it can be seen that the agreement between data and the model is reasonable in both scattering levels and patterns. When the scattering geometry approaches backscattering, i.e., the azimuthal angle between incident and scattering approaches 180°, the reverberation level (RL), after the initial direct arrival, gradually increases to its maximum, and then decreases slowly generating a broad peak as shown in (a) and (b). In contrast, when the scattering geometry changes from backward to forward, the peak becomes narrower and RL decreases faster than in the backward scattering case as shown in (c) and (d). Such a scattering characteristic is believed due to the narrow horizontal beamwidth (about 6.5°) but much broader vertical beamwidth of the transducers. For backscattering, the main lobes of the transmitter and receiver are overlapped throughout the entire data acquisition time interval associated with each ping, while for forward scattering the main lobes of the transmitter and the receiver intercept only over a much shorter time period, or may even not intercept at all (in which case the data are not analyzed further).

Another aspect of the comparison is the RL of the direct arrivals (the first large event in the time series). The differences between the prediction and the actual measurements are as large as 18 dB. The error in RL caused by the errors in position data required in estimating the geometric spreading is insignificant, less than 0.4 dB. However, the RL of the direct arrival is very sensitive to the orientation information since in many cases; the signal falls on the sharp edge (large gradient) of the main lobe or even within the sidelobes. Simulations reveal that a $\pm 1°$ error in orientation of the transducer or receiver can cause an error as large as 20 dB in estimating RL of the direct arrival. Consequently, this part of the data is not used in our analysis.

The analysis and discussions given above are also applicable to all other pings of the same data set and the other data sets. Generally speaking, the model fits the backward scattering data ($\phi_s > 100°$) better than the forward scattering data. This could be explained, in part, by the beampattern effects. For the forward scattering, the main lobes of the transmit and receive beampatterns are partially overlapped, any is more sensitive to the accuracy of the orientation of the transducers than backward scattering.

Obviously, the total scattered field at any time results from the scattering from all directions. As noted earlier, it is not possible, in general, to obtain the relationship between the scattering cross section and the scattering angles as plotted in Fig. 3 since the scattering from a certain direction characterized by a particular sub-beam cannot be separated from the total scattered field. However, by taking advantage of the narrow horizontal beamwidths of the transducers, careful inspection makes it possible to determine whether the total scattered field is dominated by scattering from a certain direction. In our simulations, at any given time, when the ratio of the intensity in one dominant direction to the intensity in any other direction is greater than 15 dB, we keep the data and designate that dominant direction as the primary scattering direction; otherwise we discard the data.

In addition, to avoid large errors due to the uncertain knowledge of sidelobes, a beampattern attenuation threshold is set to assure that all selected data fall in the main lobes. When the combined beampattern attenuation of transmitter and receiver is greater than a preset threshold (15 dB used in the modeling), the data are considered ''unreliable'' and discarded. The selected data are bounded by the vertical lines, or the time windows shown in Fig. 4. The selected data are then averaged over this time window and converted to the average scattering strength of the bubble layer by the following equation:

$$\begin{aligned} \mathrm{SS}(\gamma_2,\phi_s;\alpha_2) = {} & 10\log_{10}\langle I_{\mathrm{sd}}\rangle - \mathrm{SL} + \alpha_w r_w + \alpha_s r_{\mathrm{sed}} \\ & + 20\log_{10}(r_t) + 20\log_{10}(r_{\mathrm{scat}}) \\ & - 10\log_{10} A - 10\log_{10} B_t(\theta_t,0) \\ & - 10\log_{10} B_r(\theta_r,\phi_r) \\ & - 10\log_{10}(T_{12}T_{21}), \end{aligned} \tag{13}$$

where $\langle I_{\mathrm{sd}}\rangle$ is received scattering data averaged over the selected time window, SL is the source level, and $\alpha_{w,s} = \beta_{w,s}\log_{10} e$ is the attenuation coefficient in dB/m, and A is, as before, the ensonification area. All physical properties of the water and sediments are given in Table I. All geometry parameters are those associated with the dominant scattering direction based on the actual position and orientation information. The resultant bistatic data points are plotted in Fig. 5 (plus signs) for the mean incident grazing angle $\langle\alpha_2\rangle = 18°$ with a deviation of $\pm 10°$ and scattering grazing angle θ_s varying from 10° to 20°, where the solid lines are theoretical curves computed from our bistatic model. The theoretical curves are averages over scattering grazing angle,

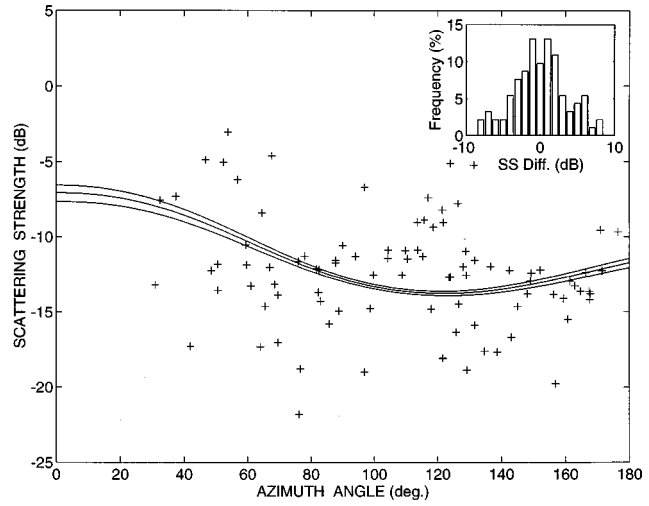


FIG. 5. Average bistatic scattering strength as a function of azimuth angle. The solid lines are the theoretical predictions given by Eq. (6) for incident grazing angle $\alpha_2 = 12° \pm 10°$ (uppermost), $18° \pm 10°$ (middle), and $24° \pm 10°$ (bottom), respectively, where we have used an exponential PDF. The experimental data (plus) are obtained for incident grazing angle $\alpha_2 = 18° \pm 10°$ and averaged over selected time window (bounded by vertical lines in Fig. 4). For both theoretical predictions and the data, the scattering grazing angle ϕ_s varies from 10° to 20°. A superimposed plot at the upper right corner is the histogram of the difference of the scattering strength between theory ($18° \pm 10°$) and the data.

θ_s $10° - 20°$ and incident grazing angles, α_2 of $12° \pm 10°$ (uppermost), $18° \pm 10°$ (middle), and $24° \pm 10°$ (bottom), respectively. The superimposed bar graph at the upper corner is the histogram of the difference between the experimental data and the theoretical curve of $\langle \alpha_2 \rangle = 18°$ with a mean $\mu = 0.0$ dB and a standard deviation $\sigma = 3.7$ dB. Further analysis reveals that about 65% of all data points deviate less than 3 dB from the theoretical prediction and about 90% less than 6 dB.

Figure 6 shows another comparison between the theory and data. The data are divided into 13 azimuth angle bands and processed by applying a median filter to the cross section

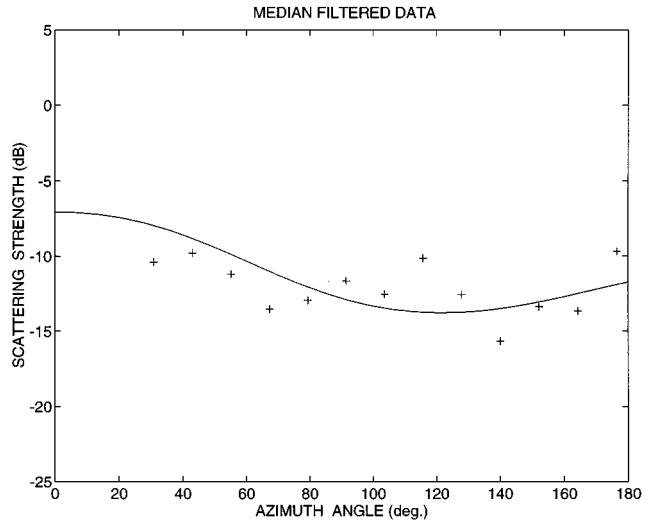


FIG. 6. Data presented in Fig. 5 are passed through a moving median filter over an azimuth angle window of 12°.

within each angle band. The model prediction and the filtered data have similar angular dependence and fluctuations about the theoretical curve are less than 4 dB.

C. Forward scattering loss and multiple scattering effects

The results presented here and in Ref. 2 demonstrate the capability of the model to quantitatively predict the measured acoustical scattering levels. One concern in the model/data comparisons is that the bubble areal density needed to obtain best agreement in the present bistatic case is 35% lower than in the backscattering case.[2] This is equivalent to a 2-dB reduction in the bistatic scattering calculated via the model. At present we have three hypotheses as to why this might be so.

One hypothesis is that the calibrations of the sonar systems used in the backscattering and bistatic experiments have a differential error of 2 dB. Error in any one calibration on the order of 1 dB cannot be ruled out.

Another possibility is that spatial variations of scattering in the region of the experiment are responsible for the difference in bubble density needed. Backscattering images in the vicinity of the tower show patchiness in the backscattering strength with variations on the order of 10 dB over scales of 10 m and a general lower level of scattering Northeast of the tower as compared to Southwest of the tower.[9] Although the bistatic scattering and backscattering data are both acquired in the vicinity of the tower, a bias in levels could remain in the data due to the differing areas of the bottom sampled.

The third hypothesis is that the different bubble density needed in backscattering and bistatic scattering is due to multiple scattering and propagation loss. Enhanced backscattering is one possible ramification of multiple scattering[10] (an enhancement of 3 dB is possible), while propagation loss due to volume scattering (single or/and multiple scattering) as the acoustic wave propagates through the bubbly sediment results in a deduction in incident intensity. For brevity, we refer this type of propagation loss as forward scattering loss. A single scattering theory excluding forward scattering loss would lead to the need for a higher bubble density in the backscattering case as compared to the bistatic case.

There are at least two other indications of multiple scattering and forward scattering loss. The first is the downward looking sonar images taken (at 12 kHz) by Lambert *et al.*[11] in the region of the experiment. In those images the bubble layer masks the signals below the layer, resulting in a dramatic reduction in the sound level below the layer.

Another (and most dramatic) indication comes from using the bistatic model to calculate the total scattered energy. The total scattered power from a unit ensonification area can be obtained by integrating the bistatic scattering cross section over all solid angles Ω,

$$P_{\text{tot}}(\alpha_2) = I_0 \int_{\Omega} \rho_b \mathscr{S}_s(\gamma_2, \phi_s; \alpha_2) d\Omega, \tag{14}$$

where I_0 is the incident intensity at the scattering layer.

If a surface scattering mechanism is assumed, the total incident power on a unit ensonified area (vertical energy flux) is $I_0 \sin(\alpha_2)$. The ratios of the total scattered power P_{tot} to the total incident power at the scattering layer are 10.5, 4.7, 2.1, and 0.8, for incident grazing angles $\alpha_2 = 5°$, 10°, 20°, and 40°, respectively. As incident angle α_2 approaches zero, the ratio approaches infinity.

The fact that the ratios are greater than unity for incident angles below 34° implies violation of the conservation of energy. Furthermore, even though the conservation of energy is not violated for the higher grazing angles, the high levels are still in violation of a single scattering assumption that relies on small scattering levels relative to the incident intensity. Further comments are included in Sec. III.

III. CONCLUSIONS

We have developed a two-layer, single scatter, bistatic model to simulate the scattering by nonspherical bubbles buried in an attenuating sediment. A bubble layer of a finite thickness is approximated by a bubble surface at a certain depth, i.e., all bubbles are on the surface. The bubbles are modeled as oblate spheroidal voids (pressure release).

When tested against data acquired in Eckernfoerde Bay, Germany, there is reasonable agreement if bubble density is reduced by 35% relative to that used in previous backscattering work for the same area. The average scattering strength exhibits a mild azimuthal dependence: Maximum scattering strength is reached in the forward direction and the minimum occurs when the transmitting and scattering directions are close to perpendicular with each other, while backscattering has a medium scattering level. The total fluctuation is about 6 dB.

However, even though the single scattering theory seems quantitatively successful for the Eckernfoerde data in some aspects, further consideration of the amount of energy scattered indicates that it cannot be the whole story. Section II C implies that for the Eckernfoerde site multiple scattering and forward scattering loss must play a role at least for some range of grazing angles in order to avoid a violation of the conservation of energy. A more severe criterion is set by the data of Lambert *et al.*[11] that indicates at least qualitatively that the two scattering mechanisms are probably important even at normal incidence. This implies that multiple scattering and forward scattering loss are in effect being treated phenomenologically in the present single scattering theory via reduction in bubble density, whereas what is probably happening is that the deeper bubbles in the bubble layer are seeing a reduced field. Consequently predictions for scattering back into the water column are consistent with those measured but predictions of the amount of energy in the sediment below the bubble layer violate both conservation of energy and the more qualitative results of Lambert *et al.*[11] Therefore, further modeling of the Eckernfoerde site that incorporates multiple scattering and forward scattering loss is desired.

This motivation for incorporation of multiple scattering and forward scattering loss is actually more general. Regardless of bubble densities, shallow incidence angles may lead to multiscattering effects. In a qualitative sense, this is because as the incidence angle gets shallower the pathlength through the bubble layer gets longer and the possibility of scattering from multiple bubbles and the forward scattering loss increases.

ACKNOWLEDGMENTS

This work is supported by the Coastal Benthic Boundary Layer Special Research Program, Office of Navel Research Grant No. N00014-95-1-G904. This is the Woods Hole Oceanographic Institution Contribution No. 9421.

APPENDIX A

To obtain an average bistatic scattering model of an oblate spheroid over orientation and aspect ratio, it is necessary to express the required spherical angles (Θ,Φ) in local coordinates (X,Y,Z) (modeling coordinates) in terms of (θ,ϕ) in the global coordinates (x,y,z) (observation coordinates). The local coordinate system is chosen in such a way that the Z axis coincides with the z axis, and the X axis coincides with the broadside unit normal of the oblate spheroid. As shown in Fig. A1, the incident and scattered wave vector can be expressed in global coordinates as

$$\hat{k}_i = (\cos\ \theta_i, 0, -\sin\ \theta_i), \tag{A1}$$

$$\hat{k}_s = (\cos\ \theta_s \cos\ \phi_s, \cos\ \theta_s \sin\ \phi_s, \sin\ \theta_s). \tag{A2}$$

The broadside unit normal of the oblate spheroid, $\hat{n}$, lies in the x-y plane and rotates about the z axis by an angle of ϕ_{rot}, representing different orientation of the oblate spheroid. To express the local coordinates (X,Y,Z) in terms of the global coordinates (x,y,z), we utilize the following coordinate transformations:

$$\hat{k}_x = l_1\hat{K}_x + l_2\hat{K}_y + l_3\hat{K}_z, \tag{A3}$$

$$\hat{k}_y = m_1\hat{K}_x + m_2\hat{K}_y + m_3\hat{K}_z, \tag{A4}$$

$$\hat{k}_z = n_1\hat{K}_x + n_2\hat{K}_y + n_3\hat{K}_z, \tag{A5}$$

where, again, upper case letters denote vector components in the local coordinate system. The direction cosines of the three rotated axes can be easily obtained

$$(l_1, m_1, n_1) = (\cos\ \phi_{\text{rot}}, \sin\ \phi_{\text{rot}}, 0), \tag{A6}$$

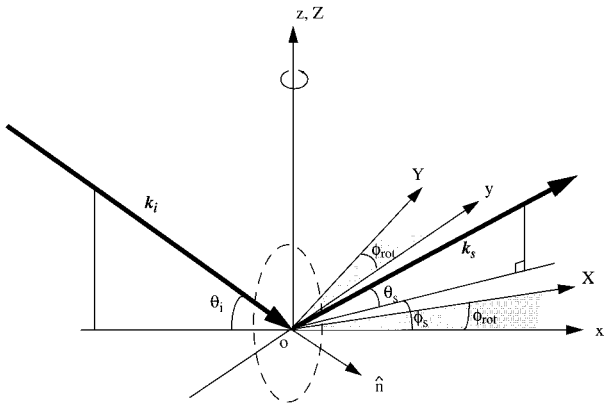


FIG. A1. Geometry of global to local coordinates transformation.

$$(l_2, m_2, n_2) = (-\sin\ \phi_{\rm rot}, \cos\ \phi_{\rm rot}, 0), \tag{A7}$$

$$(l_3, m_3, n_3) = (0,0,1), \tag{A8}$$

where $\phi_{\rm rot}$ is the rotation angle. Using these equations, we can express the unit vectors of incident and scattered waves in the local coordinates in terms of the spherical angles defined in global coordinates (x,y,z) as

$$\hat{K}_i = (\cos\ \theta_i\ \cos\ \phi_{\rm rot}, -\cos\ \theta_i\ \sin\ \phi_{\rm rot}, -\sin\ \theta_i), \tag{A9}$$

$$\hat{K}_s = (\cos\ \theta_s\ \cos(\phi_s - \phi_{\rm rot}), \cos\ \theta_s \times \sin(\phi_s - \phi_{\rm rot}), \sin\ \theta_s). \tag{A10}$$

Therefore, the two spherical angles (Θ,Φ) of the incident and scattered waves in local coordinates (X,Y,Z) are

$$\cos\ \Theta_i = \hat{K}_{ix} = \cos\ \theta_i\ \cos\ \phi_{\rm rot}, \tag{A11}$$

$$\cos\ \Theta_s = \hat{K}_{sx} = \cos\ \theta_s\ \cos(\phi_s - \phi_{\rm rot}), \tag{A12}$$

$$\tan\ \Phi_i = \frac{\hat{K}_{iz}}{\hat{K}_{iy}} = \tan\ \theta_i\ \csc\ \phi_{\rm rot}, \tag{A13}$$

$$\tan\ \Phi_s = \frac{\hat{K}_{sz}}{\hat{K}_{sy}} = \tan\ \theta_s\ \csc(\phi_s - \phi_{\rm rot}). \tag{A14}$$

It is convenient to choose a coordinate system in which the incident azimuth angle Φ_i is zero. To do so, we simply rotate the Y-Z plane about X axis by Φ_i because of the symmetry of an oblate spheroid, the new scattering azimuth angle is found to be

$$\Phi_{si} = \Phi_s - \Phi_i. \tag{A15}$$

APPENDIX B

The ensonified area is a function of the difference of the incident and scattering angles, pulse length, sound speeds of water, and sediment. For a homogeneous medium, the ensonified area at any time $t > t_{\rm min}$, where $t_{\rm min}$ is the earliest scattering arrival from the scattering layer, can be described as an elliptical ring, the outer and inner ellipses correspond two curves of equal arrival times at $t-\tau$ and t, respectively.[4] However, for a two-layer scattering geometry, since refractions across the two-layer interface greatly complicate the geometry, the ensonified area, in general, cannot be determined analytically and can only be achieved numerically. A plan view of the ensonified area for a bistatic scattering geometry is depicted schematically in Fig. B1. In the figure, $\mathbf{k_i}$ and $\mathbf{k_s}$ are incident and scattering wave vectors, $\overline{a_1 a_2}$ and $\overline{b_1 b_2}$ are wavefronts of the transmitted and scattered waves, respectively. $\overline{a_1 a_2} = x_1 \Phi_{tb}$, where x_1 is transmit horizontal distances given by Eq. (1), Φ_{tb} is the sub-beam beamwidth of the transmitter (see Sec. I A). Interception point O corresponds to the earliest arrival. The difference of arrival time between O and any point in the plane (scattering or bubble layer) is

$$\Delta_t = \frac{1}{c_2}\left(\overline{PR}/\cos\ \alpha_2 + \overline{RQ}/\cos\ \gamma_2\right), \tag{B1}$$

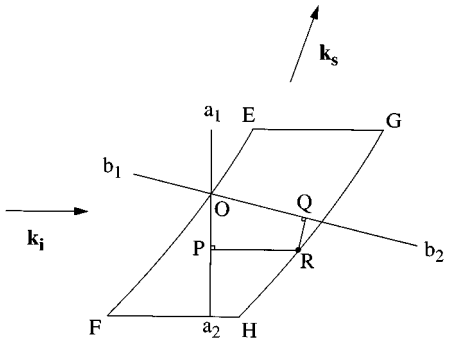


FIG. B1. Schematic description of the ensonified area for a bistatic scattering geometry.

where incident and scattering grazing angles α_2 and γ_2 can be obtained numerically. Note that $\overline{PR}$ is positive when point R is on the right of $\overline{a_1 a_2}$ and negative on the left. Positive value corresponds to a time delay while negative value corresponds to a time advance. Similarly, $\overline{QR}$ is positive when R is below $\overline{b_1 b_2}$ and negative when it is above $\overline{b_1 b_2}$.

Two arcs $\widehat{EF}$ and $\widehat{GH}$ represent two equal arrival time lines for $t-\tau$ and t, respectively. They can be obtained by setting $\Delta_t = 0$ and $\Delta_t = \tau$ in Eq. (B1), where τ is the pulse length. The ensonified area at time t for the sub-beam considered is bounded by the two equal time lines and the two segments $\overline{EG}$ and $\overline{FH}$.

[1] D. Tang, G. Jin, D. R. Jackson, and K. L. Williams, ''Analysis of high-frequency bottom and subbottom backscattering for two distinct shallow water environments,'' J. Acoust. Soc. Am. **96**, 2930–2936 (1994).

[2] D. Tang, ''Modeling high-frequency acoustic backscattering from gas voids,'' Geomarine Lett. **16**, 261–265 (1996).

[3] D. R. Jackson, ''Models for scattering from the seabed,'' Proc. Inst. Acoust. **16**, 161–169 (1994).

[4] K. L. Williams and D. R. Jackson, ''Monostatic and bistatic bottom scattering: Recent modeling and experiments,'' Proceedings of OCEANS '94, Vol. II, 255–260 (1994).

[5] A. L. Anderson and F. Abegg, ''Measurement of gas bubble concentration at Eckernfoerde Bay, Germany,'' EOS Trans. Am. Geophys. Union **75**, 159 (1994).

[6] F. Abegg, A. Anderson, L. Buzi, A. P. Lyons, and T. H. Orsi, ''Free methane concentration and bubble characteristics in Eckernfoerde Bay, Germany,'' in *Proceedings of the Gassy Mud Workshop*, edited by T. F. Wever held at the FWG, Kiel, Germany, 1994, pp. 84–89.

[7] D. T. DiPerna and T. K. Stanton, ''Sound scattering by cylinders of nonspherical cross section: A conformal mapping approach,'' J. Acoust. Soc. Am. **96**, 3064–3079 (1994). Details on the extension of cylinders to spheroids can be obtained through private communication with D. T. DiPerna.

[8] T. H. Orsi, A. L. Anderson, and A. P. Lyons, ''X-ray tomographic analysis of sediment macrostructure and physical property variability in Eckernfoerde Bay sediments, Western Baltic Sea,'' Geomarine Lett. **16**, 232–239 (1996).

[9] D. R. Jackson, K. L. Williams, and K. B. Briggs, ''High-frequency acoustic observations of benthic spatial and temporal variability,'' Geomarine Lett. **16**, 212–218 (1996).

[10] A. Ishimaru, ''Backscattering enhancement: From radar cross-sections to electron and light localizations to rough surface scattering,'' IEEE Trans. Antennas Propag. **33**, 7–11 (1991).

[11] D. N. Lambert, D. J. Walter, and J. A. Hawkins, ''Delineation of shallow, subbottom gas concentrations using a narrow beam, high frequency acoustic system,'' Proceedings of the workshop on Modeling Methane-Rich sediments of Eckernfoerde Bay, Eckernfoerde, 26–30 June 1995.

Reconstruction of the velocity and density in a stratified acoustic half-space using a short-pulse point source

Jiasu Cao and Sailing He
Department of Electromagnetic Theory, Royal Institute of Technology, S-100 44 Stockholm, Sweden

(Received 21 August 1995; revised 1 July 1996; accepted 1 April 1997)

The reconstruction of the velocity and density in a stratified half-space using the reflection data generated by a short-pulse point source is considered. Transverse spatial Hankel transform is applied to the transient reflected and incident pressures on the surface. The reflection kernel in the Hankel-transformed space is obtained through a deconvolution. The time-domain Green function approach based on the wave splitting is then used to reconstruct the parameters. This time-domain approach to the inverse problem is exact and noniterative. Numerical results for some synthetical examples are presented to illustrate the Hankel transform, deconvolution, and the reconstruction. The finite-difference time-domain method is used to generate the transient reflection data in the physical space in a numerical simulation. The reconstruction using obliquely incident plane waves is also considered in an Appendix.
[S0001-4966(97)00808-4]

PACS numbers: 43.30.Pc, 43.30.Ky, 43.20.Fn [JHM]

INTRODUCTION

A useful technique for the time-domain inverse problems is based on the method of wave splitting, i.e., the decomposition of the total field into two components which propagate in opposite directions. In a homogeneous region the definition of the split fields reflects the factorization of the wave equation, and projects out the down- and up-going parts of the solution. The time-domain wave splitting was first combined with the invariant imbedding idea for one-dimensional inverse scattering problems, cf., e.g., Refs. 1–5. Given an inhomogeneous medium and a splitting, one can define associated scattering operators. Invariant imbedding techniques then allow one to obtain partial differential equations (PDEs) for these scattering operators. The imbedding equation for the reflection operator is nonlinear. In Ref. 6, a Green function approach based on wave splitting has been introduced to map the incident field to the internal split fields in a stratified medium for a direct problem. In contradistinction to the invariant imbedding method, the system of equations for the Green functions is linear and is suitable for parallel processing, cf., e.g., Ref. 7. The wave-splitting approach has also been applied to the direct and inverse scattering problems in the frequency domain (see, e.g., Refs. 8–10). The wave-splitting approach has been shown to be an efficient method for direct and inverse problems concerning wave propagation in linear media.

In the present paper, we apply the time-domain Green function approach to the acoustic inverse problem for a short-pulse point source over a stratified half-space. Most numerical results obtained from the wave-splitting and imbedding or Green's function approaches, which can be found in the literature, concern the inverse problems in which some scattering kernel (e.g., the reflection kernel) is taken as given data. In the present work the ill-posed step (deconvolution) between the measured fields and the reflection kernel is included in the numerical implementation. In fact, it has been shown that some inverse problems relating to plane-stratified media in which the reflection kernel is given data, are well-posed (see Ref. 11). The wave-splitting Green function approach provides a solution to both the direct and (some) inverse problems, and the forward data can be (and usually is) used in the inverse problem. However, it is desirable to compute the forward data by a different approach and therefore we also use a finite-difference method to get the forward data. Comparisons are made between the results obtained from the two ways of getting the forward data. We also exploit the possibility to use a previously derived explicit and exact result concerning the reflection kernel for an exponentially stratified half-space by the authors[12] as a check. Inverse acoustic problems have been considered by others using different methods (see, e.g., Refs. 13–17). In the present paper, we apply a transverse spatial Hankel transform to the transient reflected and incident pressures on the surface. The reflection kernel in the Hankel-transformed space is obtained through a deconvolution. The time-domain Green function approach is then used to reconstruct the parameters with the reflection kernel known as the input. Partial differential equations for the Green functions are given together with the initial and boundary conditions. The present time-domain approach to the inverse problem is exact and noniterative. Numerical results are presented to illustrate the Hankel transform, deconvolution, and reconstruction for a synthetical example, in which the transient reflection data in the physical space is generated by the finite-difference time-domain (FDTD) method. Numerical results for a simultaneous reconstruction of the velocity and density using two values of the Hankel-transform parameter are also presented. The case of the obliquely incident plane waves is also considered in the Appendix, where the numerical results for the reconstruction using two incident angles are presented. The present approach has also recently been developed for the inverse problems for three-dimensional inhomogeneous media; see, e.g., Refs. 18 and 19.

I. PROBLEM FORMULATION AND WAVE-SPLITTING

Consider transient wave propagation in an inhomogeneous acoustic medium which is described by the following partial differential equation for the pressure p:[20]

$$\frac{1}{c^2}\frac{\partial^2 p}{\partial t^2}-\rho\nabla\cdot\left(\frac{1}{\rho}\nabla p\right)=0, \tag{1}$$

where ρ is the fluid density and c is the velocity of the acoustic wave in the medium. Note that the above acoustic wave equation does not include any loss mechanism. However, the present approach can be easily generalized to the lossy case.[21] In the present paper we consider the case of a point source over a stratified acoustic half-space. The velocity and density are functions of the depth z in the half-space $z\geqslant 0$. The upper half-space $z<0$ is homogeneous and the point source is located at the position $(0,0,z_0)$ with $z_0<0$. For simplicity of analysis, it is assumed that the velocity and the density are continuous everywhere, and the phase velocity is furthermore continuously differentiable for $z>0$ [the discontinuous case can be exactly treated in an analogous way with some modifications of the scattering coefficients, cf., e.g., Ref. 22; steep slopes are used to approximate jump discontinuities of the parameters in the numerical treatment, cf. Figs. 6(b) and A1(b)]. Noticing the axial symmetry of the scattering configuration, we obtain the following equations for $p=p(r,z,t)$ from Eq. (1) in the cylindrical coordinates (r,ϕ,z),

$$\frac{1}{c^2}\frac{\partial^2 p}{\partial t^2}-\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial p}{\partial r}\right)-\frac{\partial^2 p}{\partial z^2}+\frac{d}{dz}(\ln\rho)\frac{\partial p}{\partial z}=0, \quad z>0, \tag{2}$$

$$\frac{1}{c_0^2}\frac{\partial^2 p}{\partial t^2}-\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial p}{\partial r}\right)-\frac{\partial^2 p}{\partial z^2}=0, \quad z<0,\quad (z-z_0)^2+r^2>0, \tag{3}$$

where $c_0=c(0)$ is the wave speed in the homogeneous upper half-space. It is assumed that the incident wave will not reach the interface $z=0$ until the time $t=0$, i.e., we have the initial conditions

$$p(r,z,0)=\frac{\partial p}{\partial t}(r,z,0)=0, \quad \text{for } z>0. \tag{4}$$

In view of the axial symmetry of the scattering problem, the following Hankel transform is introduced

$$\overline{p}(z,t;\kappa)=\int_0^\infty p(r,z,t)J_0(\kappa r)r\,dr, \tag{5}$$

$$p(r,z,t)=\int_0^\infty \overline{p}(z,t;\kappa)J_0(\kappa r)\kappa\,d\kappa, \tag{6}$$

where J_0 is the zeroth-order Bessel function.

Thus from the transformation of Eqs. (2)–(4) we obtain

$$\frac{\partial^2\overline{p}}{\partial z^2}-\frac{1}{c^2}\frac{\partial^2\overline{p}}{\partial t^2}-\kappa^2\overline{p}-\frac{d}{dz}(\ln\rho)\frac{\partial\overline{p}}{\partial z}=0, \quad z>0, \tag{7}$$

$$\frac{\partial^2\overline{p}}{\partial z^2}-\frac{1}{c_0^2}\frac{\partial^2\overline{p}}{\partial t^2}-\kappa^2\overline{p}=0, \quad z<0, \quad (z-z_0)^2+r^2>0, \tag{8}$$

$$\overline{p}(z,0;\kappa)=\frac{\partial\overline{p}}{\partial t}(z,0;\kappa)=0, \quad \text{for } z>0. \tag{9}$$

In the Hankel-transformed space, a wave splitting that splits the total pressure $\overline{p}$ into a down-going pressure $\overline{p^+}$ and an up-going pressure $\overline{p^-}$ is defined as[21] [cf. Ref. 3 for the splitting in the (x,y,z) space]

$$\begin{bmatrix}\overline{p^+}\\ \overline{p^-}\end{bmatrix}=\frac{1}{2}\begin{bmatrix}1 & -K\\ 1 & K\end{bmatrix}\begin{bmatrix}\overline{p}\\ \dfrac{\partial\overline{p}}{\partial z}\end{bmatrix}\equiv T\begin{bmatrix}\overline{p}\\ \dfrac{\partial\overline{p}}{\partial z}\end{bmatrix}, \tag{10}$$

where the splitting operator K and its inverse K^{-1} satisfy the following equation:

$$K^{-1}=\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2}+\kappa^2\right)K. \tag{11}$$

The wave splitting in a homogeneous region is obtained as a consequence of the obvious factorization of the wave operator, i.e.,

$$\frac{\partial^2\overline{p}}{\partial z^2}-\frac{1}{c^2}\frac{\partial^2\overline{p}}{\partial t^2}-\kappa^2\overline{p}=\left[\frac{\partial}{\partial z}+K^{-1}\right]\left[\frac{\partial}{\partial z}-K^{-1}\right]\overline{p}.$$

Thus $\overline{p^+}$ and $\overline{p^-}$ have physical interpretation as the transformed incident pressure and the transformed reflected pressure, respectively, in the region $z_0<z\leqslant 0$.

Equation (7) can be written in a matrix form as

$$\partial_z\begin{bmatrix}\overline{p}\\ \dfrac{\partial\overline{p}}{\partial z}\end{bmatrix}=\begin{bmatrix}0 & 1\\ \dfrac{1}{c^2}\dfrac{\partial^2}{\partial t^2}+\kappa^2 & \dfrac{d}{dz}(\ln\rho)\end{bmatrix}\begin{bmatrix}\overline{p}\\ \dfrac{\partial\overline{p}}{\partial z}\end{bmatrix}\equiv D\begin{bmatrix}\overline{p}\\ \dfrac{\partial\overline{p}}{\partial z}\end{bmatrix}. \tag{12}$$

Differentiating Eq. (10) with respect to z and using Eq. (12), yields

$$\partial_z\begin{bmatrix}\overline{p^+}\\ \overline{p^-}\end{bmatrix}=TDT^{-1}\begin{bmatrix}\overline{p^+}\\ \overline{p^-}\end{bmatrix}+\frac{\partial T}{\partial z}T^{-1}\begin{bmatrix}\overline{p^+}\\ \overline{p^-}\end{bmatrix}\equiv\begin{bmatrix}\alpha & \beta\\ \gamma & \delta\end{bmatrix}\begin{bmatrix}\overline{p^+}\\ \overline{p^-}\end{bmatrix}, \tag{13}$$

where

$$\alpha=-K^{-1}+\frac{\rho_z}{2\rho}+\frac{c_z}{2c}(1-\kappa^2K^2), \tag{14}$$

$$\beta=\gamma=-\frac{\rho_z}{2\rho}-\frac{c_z}{2c}(1-\kappa^2K^2), \tag{15}$$

$$\delta=K^{-1}+\frac{\rho_z}{2\rho}+\frac{c_z}{2c}(1-\kappa^2K^2), \tag{16}$$

where $c_z=(d/dz)c$, $\rho_z=(d/dz)\rho$. Note that in the above derivation, the following relation is used:[21]

$$\frac{\partial K}{\partial z}K^{-1}=-K\frac{\partial K^{-1}}{\partial z}=\frac{c_z}{c}(1-\kappa^2K^2).$$

Equation (13) gives the dynamics for the split pressures in the Hankel-transformed space.

Furthermore it has been proved that the operators K^{-1} and K^2 have the following explicit representations:[21]

$$K^{-1}f(z,t;\kappa)=\frac{1}{c(z)}\frac{\partial}{\partial t}f(z,t;\kappa)+A(z,t;\kappa)*f(z,t;\kappa), \tag{17}$$

$$K^2 f(z,t;\kappa)=C(z,t;\kappa)*f(z,t;\kappa), \tag{18}$$

where

$$A(z,t;\kappa)=\kappa J_1(\kappa c(z)t)/t, \tag{19}$$

$$C(z,t;\kappa)=\frac{c(z)}{\kappa}\sin(\kappa c(z)t). \tag{20}$$

The $*$ in Eqs. (17) and (18) denotes the time convolution integral, i.e.,

$$f(z,t;\kappa)*g(z,t;\kappa)=\int_0^t f(z,t-t';\kappa)g(z,t';\kappa)dt'. \tag{21}$$

II. GREEN FUNCTION APPROACH TO THE INVERSE PROBLEM

The Green functions $G^{\pm}(z,t;\kappa)$ are defined as the convolution kernels of the mappings from the incident pressure at $z=0$ to the internal split pressures in the Hankel-transformed space according to

$$\begin{bmatrix}\overline{p^+}\\ \overline{p^-}\end{bmatrix}(z,t+\tau(z);\kappa)=a(z)\begin{bmatrix}\overline{p^+}(0,t;\kappa)\\ 0\end{bmatrix}+\begin{bmatrix}\int_0^t G^+(z,t';\kappa)\overline{p^+}(0,t-t';\kappa)dt'\\ \int_0^t G^-(z,t';\kappa)\overline{p^+}(0,t-t';\kappa)dt'\end{bmatrix}, \tag{22}$$

where $\tau(z)$ corresponds to the vertical travel time from 0 to z, i.e.,

$$\tau(z)=\int_0^z\frac{dz'}{c(z')}, \tag{23}$$

and $a(z)$ is the attenuation factor

$$a(z)=\sqrt{\frac{\rho(z)c(z)}{\rho(0)c(0)}}. \tag{24}$$

The appearance of the time convolution integrals in definitions (22) is due to the fact that the response of the medium is causal and invariant under time translation.

Differentiating Eq. (22) with respect to z, we obtain the following equation:

$$\frac{\partial}{\partial z}\begin{bmatrix}\overline{p^+}\\ \overline{p^-}\end{bmatrix}(z,t+\tau)+\frac{1}{c}\frac{\partial}{\partial t}\begin{bmatrix}\overline{p^+}\\ \overline{p^-}\end{bmatrix}(z,t+\tau)=\left(\frac{da}{dz}\right)\begin{bmatrix}\overline{p^+}(0,t;\kappa)\\ 0\end{bmatrix}+\begin{bmatrix}\frac{\partial G^+}{\partial z}(z,t;\kappa)*\overline{p^+}(0,t;\kappa)\\ \frac{\partial G^-}{\partial z}(z,t;\kappa)*\overline{p^+}(0,t;\kappa)\end{bmatrix}.$$

Substituting the dynamics system (12) into the above equation, we obtain the following partial differential equations (PDEs) for the Green functions (see Ref. 21 for a detailed derivation in a similar case):

$$\frac{\partial}{\partial z}G^+=-aA-\frac{ac_z}{2c}\kappa^2 C-A*G^+ +\frac{1}{2}\left[\frac{c_z}{c}+\frac{\rho_z}{\rho}\right]\times(G^+-G^-)-\frac{c_z}{2c}\kappa^2 C*(G^+-G^-), \tag{25}$$

$$\frac{\partial}{\partial z}G^-=\frac{2}{c}\frac{\partial}{\partial t}G^- +\frac{ac_z}{2c}\kappa^2 C+A*G^- -\frac{1}{2}\left[\frac{c_z}{c}+\frac{\rho_z}{\rho}\right](G^+-G^-)+\frac{c_z}{2c}\kappa^2 C*(G^+-G^-), \tag{26}$$

together with the following initial condition for G^-

$$G^-(z,0;\kappa)=\frac{1}{4}a\left(c_z+\frac{c\rho_z}{\rho}\right). \tag{27}$$

From definition (22) for the Green functions, we have the following boundary conditions:

$$G^+(0,t;\kappa)=0, \tag{28}$$

$$G^-(0,t;\kappa)=R(t;\kappa), \tag{29}$$

where $R(t;\kappa)$ is the reflection kernel [i.e., the reflection due to an impulsive $\overline{p^I}(0,t;\kappa)$] defined as[21]

$$\overline{p^r}(0,t;\kappa)=\int_0^t R(t';\kappa)\overline{p^I}(0,t-t';\kappa)dt', \tag{30}$$

where $\overline{p^r}(0,t;\kappa)$ and $\overline{p^I}(0,t;\kappa)$ are the reflected and incident pressures, respectively, at the surface $z=0$ in the Hankel-transformed space.

In the inverse problem, the reflection kernel $R(t;\kappa)$ is obtained from the measured transient reflection on the surface; see Sec. III B.

If the reflection kernel $R(t;\kappa)$ is known for a fixed value of κ, we can propagate the known boundary values [cf. Eqs. (28) and (29)] to the initial values using the PDEs for the Green functions. The velocity $c(z)$ or the density $\rho(z)$ can then be reconstructed according to initial conditions (27). The numerical algorithm and results for the reconstruction will be given in Sec. IV. The input data for the reconstruction is the reflection kernel $R(t;\kappa)$ with a fixed value of κ. In the next section we will show how to obtain $R(t;\kappa)$.

III. OBTAINING THE REFLECTION KERNEL

In this section, we compute the reflected pressure in physical space with the finite-difference time-domain (FDTD) method, and then take the transverse Hankel transform (with a fixed value of κ) for the reflected pressure at the

surface, as well as for the incident pressure. The reflection kernel $R(t;\kappa)$ is then obtained from a deconvolution according to Eq. (30).

A. Generating reflection data in physical space using the FDTD method

We use the FDTD method to generate the reflected pressure for a short-pulse point source [with a time-varying function $f(t)$] over a half-space with a stratified density (the velocity is taken to be a constant c_0 in this numerical example in order to check with a previously derived explicit and exact result, cf. Ref. 12).

In the xyz coordinate system, the incident pressure $p^i(x,y,z,t)$ satisfies the following equation:

$$\frac{1}{c_0^2}\frac{\partial^2 p^i}{\partial t^2}-\nabla^2 p^i=\delta(\mathbf{r}-\mathbf{r}_0)f(t), \tag{31}$$

where $\mathbf{r}_0=(0,0,z_0)$. Obviously the solution to the above equation is

$$p^i=\frac{f(t-|\mathbf{r}-\mathbf{r}_0|/c_0)}{4\pi|\mathbf{r}-\mathbf{r}_0|}. \tag{32}$$

The total pressure $p(x,y,z,t)$ satisfies the following equation:

$$\frac{1}{c_0^2}\frac{\partial^2 p}{\partial t^2}-\rho\nabla\cdot\left(\frac{1}{\rho}\nabla p\right)=\delta(\mathbf{r}-\mathbf{r}_0)f(t). \tag{33}$$

From Eqs. (31) and (33), we obtain the following equation for the scattered pressure $p^s(x,y,z,t)\equiv p(x,y,z,t)-p^i(x,y,z,t)$:

$$\frac{1}{c_0^2}\frac{\partial^2 p^s}{\partial t^2}-\nabla^2 p^s+\frac{\rho_z}{\rho}\frac{\partial}{\partial z}p^s=-\frac{\rho_z}{\rho}\frac{\partial}{\partial z}p^i. \tag{34}$$

Using central difference formulas to approximate the above partial differential equation, one obtains the following discretized equation:

$$\begin{aligned}p^s(i,j,k,n+1)=&\,2p^s(i,j,k,n)-p^s(i,j,k,n-1)+(c_0\Delta t/\Delta z)^2\{[p^s(i+1,j,k,n)-2p^s(i,j,k,n)+p^s(i-1,j,k,n)]\\&+[p^s(i,j+1,k,n)-2p^s(i,j,k,n)+p^s(i,j-1,k,n)]+[p^s(i,j,k+1,n)-2p^s(i,j,k,n)+p^s(i,j,k-1,n)]\\&-\tfrac{1}{2}\Delta z(\rho_z/\rho)(i,j,k)[p^s(i,j,k+1,n)-p^s(i,j,k-1,n)+p^i(i,j,k+1,n)-p^i(i,j,k-1,n)]\},\end{aligned} \tag{35}$$

where $p^s(i,j,k,n)=p^s(i\Delta x,j\Delta y,k\Delta z,n\Delta t)$ with $\Delta x=\Delta y=\Delta z$, etc. The initial values are

$$p^s(i,j,k,-1)=p^s(i,j,k,0)=0, \tag{36}$$

due to the fact that there is no scattering before the time $t=0$ (the incident wave has not reached the interface $z=0$ yet). This algorithm is stable when (see, e.g., Refs. 23 and 24)

$$\Delta t\leqslant\frac{\Delta z}{\sqrt{3}c_0}. \tag{37}$$

As a numerical example, we choose the short pulse to have the following time dependence:

$$f(t)=\begin{cases}\sin\left(\dfrac{t+|z_0|/c_0}{t_0}\,\pi\right),\\ \text{when}\ \ -|z_0|/c_0\leqslant t\leqslant -|z_0|/c_0+t_0,\\ 0,\ \text{otherwise.}\end{cases} \tag{38}$$

Thus it has nonzero values only within a time width t_0, and reaches the interface $z=0$ at the time $t=0$ [note that the position of the point source is $(0,0,z_0)$]. The density profile used in this numerical example is

$$\rho(z)=\rho_0\frac{e^{-2\beta z}(1+\alpha)^2}{(e^{-2\beta z}+\alpha)^2},\quad z>0, \tag{39}$$

where $\rho_0=1$, $\beta=1$, and $\alpha=0.1$. The velocity is $c=c_0=1$. We choose $z_0=-0.5$ for the position of the point source, and $t_0=0.4$ in Eq. (38) for the width of the short pulse in this example. Numerical computation domains have to be finite. In this example, the boundary surfaces of the computation domain are set to be far enough away that putting p^s to zero at the boundary surfaces will have no effect on p^s at the receiving points and during the time periods of interest [the source point can be in the computation domain in this example since p^i is zero (nonsingular) at the source point for $t>0$; note that $t_0<|z_0|/c_0$ in this example]. On the interface $z=0$, p^s is identical to the reflected pressure p^r. We choose $\Delta z=\frac{1}{32}$, $\Delta t=0.2(\Delta z)/c_0$ due to the limited storage space of our computer. The numerical results for the reflected pressure p^r along a line on the interface $z=0$ are plotted in Fig. 1.

B. Obtaining the reflection kernel from the reflection data on the surface

1. Hankel transform

On the surface $z=0$, we take a Hankel transform (5) (for a fixed value of κ) of the reflected pressure p^r calculated by the FDTD method in the previous subsection, as well as of the incident pressure p^i [cf. Eq. (32)]. Since we use a time-domain method and the pressure $p(r,z,t)$ is nonzero only within a finite region $r<r_1$ on the surface $z=0$ at any finite time $t=t_1$, the transverse spatial Hankel transform (5) does not have any convergence problem. The numerical results are plotted in Fig. 2, where the solid and dashed lines are for the Hankel-transformed (with $\kappa=1$) reflected and incident pressures, respectively, at the surface $z=0$.

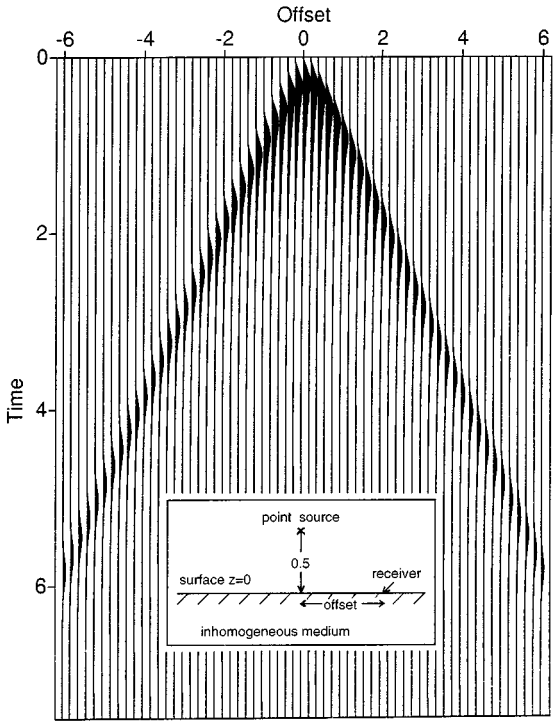


FIG. 1. The transient reflected pressure along a line on the surface $z=0$, for the density profile given by Eq. (39) with $\rho_0=1$, $\beta=1$, and $\alpha=0.1$ (see the solid line in Fig. 5; the velocity $c=c_0=1$). The short-pulse point source is described by Eqs. (32) and (38) with $z_0=-0.5$, $t_0=0.4$. The numerical results are obtained through the FDTD method.

2. Deconvolution

After the Hankel transformation of the reflected and incident pressures on the surface $z=0$, the reflection kernel $R(t;\kappa)$ is now obtained from a deconvolution according to Eq. (30). The deconvolution program used here is based on the fast Fourier transform of (30). In a deconvolution process we have to take a cutoff frequency ω_{cutoff}, since the high-frequency components are mainly noise introduced from the measurement (in our numerical simulation it is the computational errors of the finite-difference method) and the Hankel transform. The deconvolution is an ill-posed problem and a regularization filter is necessary to yield a stable and physically consistent result. Let $\tilde{R}(\omega;\kappa)$ be the spectrum of $R(t;\kappa)$; we then use the following formula:

$$\tilde{R}(\omega;\kappa)=\frac{\mathcal{F}[\overline{p^r}(0,t;\kappa)]}{\mathcal{F}[\overline{p^i}(0,t;\kappa)]} \approx \frac{\{\mathcal{F}[\overline{p^r}(0,t;\kappa)]\}\{\mathcal{F}[\overline{p^i}(0,t;\kappa)]\}^*}{\|\mathcal{F}[\overline{p^i}(0,t;\kappa)]\|^2+\lambda\omega^4}, \tag{40}$$

where λ is a filtering constant (see Ref. 25 for more details), and the superscript * denotes the complex conjugate. $R(t;\kappa)$ is then obtained from the inverse Fourier transform of the spectrum $\tilde{R}(\omega;\kappa)$.

The deconvolution result is plotted in Figs. 3 and 4. The spectra $\mathcal{F}[\overline{p^r}(0,t;\kappa=1)]$ and $\mathcal{F}[\overline{p^i}(0,t;\kappa=1)]$ (their amplitudes) are plotted in Fig. 3(a), and the spectrum $\mathcal{F}[R(t;\kappa=1)]$ obtained through Eq. (40) is shown in Fig. 3(b). The reflection kernel $R(t;\kappa=1)$ obtained from a deconvolution with the cutoff frequency $\omega_{\text{cutoff}}=4.5$ is presented in Fig. 4 by the solid line. The exact solution for $R(t;\kappa)$ given by Cao and He[12] [cf. Eq. (41) in the next section] with $\kappa=1$ is also plotted in Fig. 4 by the dotted line. As seen from Fig. 4, the

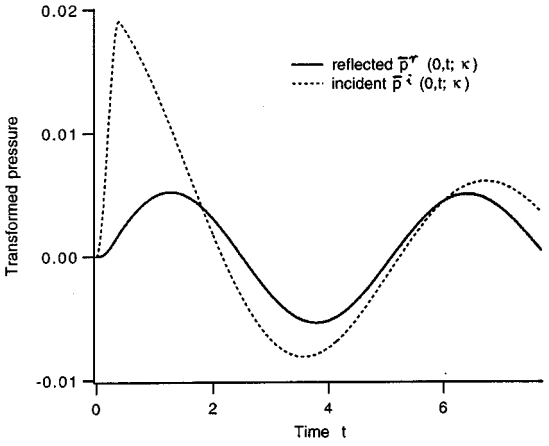


FIG. 2. The corresponding Hankel-transformed reflected and incident pressures at $z=0$. The value for the Hankel-transform parameter is $\kappa=1$.

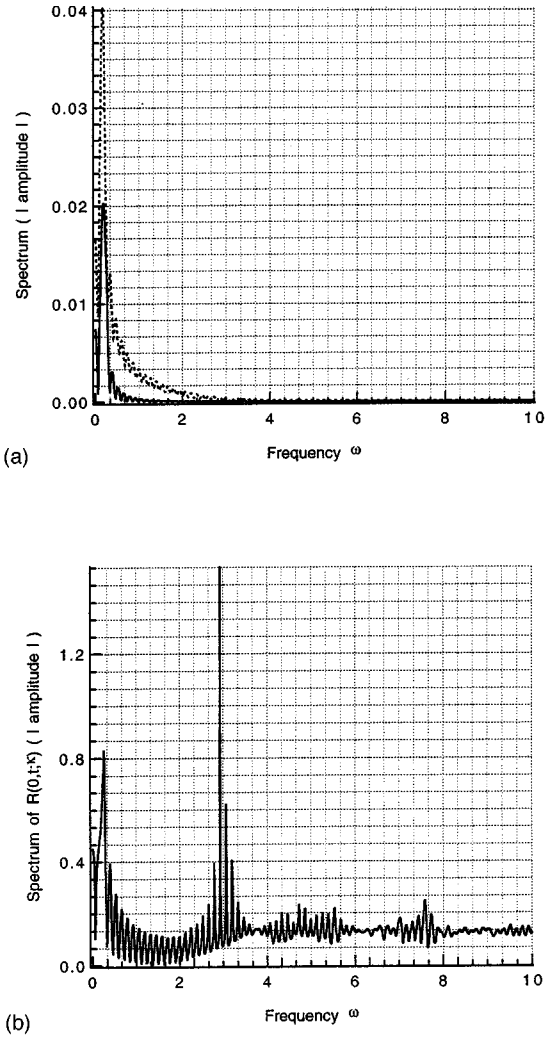


FIG. 3. (a) The corresponding spectra $\mathcal{F}[\overline{p^r}(0,t;\kappa=1)]$ (the solid line) and $\mathcal{F}[\overline{p^i}(0,t;\kappa=1)]$ (the dotted line) for the example used in Fig. 1. (b) The corresponding spectrum $\mathcal{F}[R(t;\kappa=1)]$.

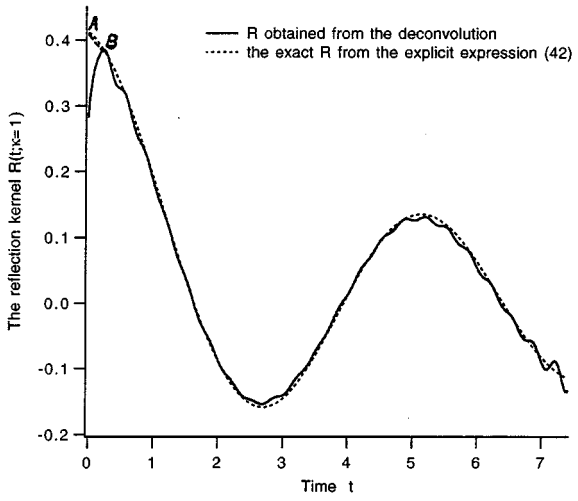


FIG. 4. The reflection kernel $R(t;\kappa=1)$ (the solid line) obtained from a deconvolution with the cutoff frequency $\omega_{\text{cutoff}}=4.5$, for the example used in Fig. 1. The dotted line is from the exact expression (41). The dashed line segment AB is the correction to the deconvolution for small t (using the prior knowledge of R at $t=0$).

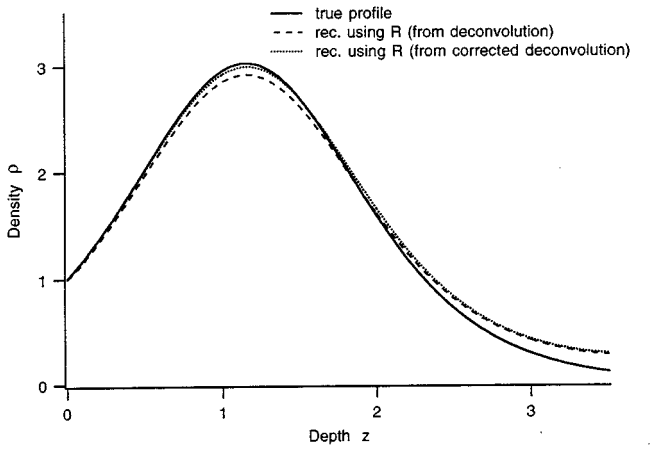


FIG. 5. Reconstruction of the density using the reflection kernel obtained from the deconvolution. The solid line is the true profile. The dashed line is the reconstruction using $R(t;\kappa=1)$ directly obtained from the deconvolution (the solid line in Fig. 4). The dotted line is the reconstruction using $R(t;\kappa=1)$ obtained from the corrected deconvolution (the dashed line segment AB and the remaining part of the solid curve in Fig. 4).

reflection kernel $R(t;\kappa)$ obtained from the deconvolution is consistent with the exact $R(t;\kappa)$ except for small t. The main reason for the error at small t is due to the use of the cutoff frequency in the deconvolution process (we only use the information with the frequencies lower than ω_{cutoff} in the spectra of the reflected and incident pressures). We note that

$$R(t;\kappa)|_{t=0}=\int_{-\infty}^{\infty}\tilde{R}(\omega,\kappa)d\omega.$$

Note that practically the bandwidth does not need to be infinite for the present inversion technique to be successful.

If we know the parameters at the surface $z=0$, we can obtain the correct value of $R(t=0;\kappa)$ from the initial condition (27), and then replace the beginning portion of the $R(t;\kappa)$ curve for small t with a straight short line segment connecting to a point with a reasonably reliable value of R (e.g., point B in Fig. 4), as shown by the dashed short line segment AB in Fig. 4 [this will be referred to as $R(t;\kappa)$, obtained from a corrected deconvolution, which will also be used in the reconstruction in the next section]. We can also set the surface of the inhomogeneous medium a bit lower to the plane $z=z_1>0$ so that the parameters in the region $0\leqslant z\leqslant z_1$ are known (e.g., the air), which means $R(t;\kappa)$ is known for a short period of times $0\leqslant t\leqslant t_1$. In such a way we can remove the error in $R(t;\kappa)$ (obtained from a deconvolution) for small $t\in(0,t_1)$. Nevertheless, the errors in $R(t;\kappa)$ for small t are not critical for the Green function approach to the reconstruction, as we will see in the next section.

In a real experiment, we measure the total transient pressure along a line on the surface $z=0$ for a finite period of time, and obtain the reflected pressure by subtracting the incident pressure (generated by a point source with short time-varying pulse). We choose a specific value of the Hankel-transform parameter κ to transform the surface (incident and reflected) data. Then we perform the deconvolution to obtain the reflection kernel $R(t;\kappa)$ for this chosen κ, as described above in this subsection. Once the reflection kernel $R(t;\kappa)$ is obtained, we can use the Green function approach described in the previous section to reconstruct the parameters in the stratified half-space.

IV. NUMERICAL RECONSTRUCTION

A. Example 1

Example 1 demonstrates reconstruction of the density using the reflection kernel obtained from the deconvolution. The reflection kernel $R(t;\kappa=1)$ (the solid line in Fig. 4) obtained from the deconvolution in the previous section is now used as the boundary value of the Green function G^- [cf. Eq. (29)] in the Green function approach to reconstruct the density $\rho(z)$. After propagating the known boundary values [cf. Eqs. (28) and (29)] to the initial values using the PDEs for the Green functions, the density $\rho(z)$ is reconstructed using the initial conditions (27) (see Ref. 21 for a more detailed description of the numerical algorithm for the Green function approach). The reconstructed density $\rho(z)$ is plotted by the dashed line in Fig. 5, where the solid line is the true profile given by Eq. (39). We note that even though $R(t;\kappa=1)$ used here is quite different from the true value for small t (cf. Fig. 4), the reconstruction is still good. This indicates that the values of the reflection kernel $R(t;\kappa)$ for small t are not critical for reconstruction in the present Green function approach. The reconstruction can be improved by using the reflection kernel $R(t;\kappa=1)$ (the dashed line segment AB and the remaining part of the solid curve in Fig. 4) obtained from a corrected deconvolution, as shown by the dotted line in Fig. 5. Note that the error of the reconstruction at the deep positions in Fig. 5 is due to the computation errors introduced in the process of FDTD, Hankel transform, and deconvolution.

As shown in Ref. 12, the reflection kernel for this special density profile (when $\alpha\neq 1$) has the following exact explicit form:

$$R(t;\kappa)=\frac{1}{a}\left[\frac{\sqrt{\kappa^2+\beta^2}}{\beta}\cdot\frac{J_1(c_0t\sqrt{\kappa^2+\beta^2})}{t}-\frac{\kappa}{\beta}\cdot\frac{J_1(c_0\kappa t)}{t}\right]+\frac{a^2-1}{a}\cdot c_0\beta f_2(t;\kappa), \tag{41}$$

where $a=1-\alpha/1+\alpha$ and

$$f_2(t;\kappa)=\frac{1}{2}f_1(t;\kappa)+\frac{1}{2a}f_1(t;\kappa)*\left[\frac{\sqrt{\kappa^2+\beta^2}}{\beta}\cdot\frac{J_1(c_0t\sqrt{\kappa^2+\beta^2})}{t}-\frac{\kappa}{\beta}\cdot\frac{J_1(c_0\kappa t)}{t}\right],$$

$$f_1(t;\kappa)=J_0(c_0\kappa t)-\frac{(1-a^2)c_0\beta}{2a}\times\int_0^t J_0(c_0\kappa\sqrt{t^2-t'^2})e^{-[(1-a^2)c_0\beta/2a]t'}\,dt'.$$

The exact $R(t;\kappa)$ given by Eq. (41) for $\kappa=1$ is plotted by the dotted line in Fig. 4. Synthetically, we can also generate $R(t;\kappa)$ by propagating (i.e., integrating along the characteristics) the initial values of the Green functions [cf. Eq. (27); the initial values for G^+ can be obtained by integration of Eq. (25) along $+z$ direction] to the boundary values using the PDEs for the Green functions [note that the boundary value of G^- gives R, cf. Eq. (29)]. The numerical result for $R(t;\kappa)$ obtained by such an integration of the PDEs for $G^{\pm}$ is identical to the explicit exact solution (dotted line in Fig. 4) on the scale of Fig. 4 when the number of the discrelization points in z is greater than 128. Since the reflection kernel obtained through a process of FDTD computation, Hankel transform, and deconvolution is consistent with the explicit exact solution (except for small t; cf. Fig. 4), the reflection kernel obtained through the integration of the PDEs for $G^{\pm}$ along the characteristics is also consistent with the one obtained through a process of FDTD computation, Hankel transform, and deconvolution (except for small t). Therefore, in the numerical example of a simultaneous reconstruction given below, instead of through a process of FDTD computation, Hankel transform, and deconvolution, we generate the reflection kernel $R(t;\kappa)$ through integration of the PDEs for $G^{\pm}$ along the characteristics [this certainly gives higher accuracy of the reflection kernel $R(t;\kappa)$ for use as an input for the inverse problem].

B. Example 2

Example 2 shows simultaneous reconstruction using two different values of κ. In this numerical example, we will use the reflection kernel $R(t;\kappa)$ for two different values of κ to reconstruct the velocity $c(z)$ and the density $\rho(z)$ simultaneously. Theoretical conditions for simultaneous reconstruction in a similar case have been reported by He and Karlsson,[21] and the relation between the Green functions and the imbedding reflection kernel $R(z,t;\kappa)$ was used there. In the present paper we introduce a function $R_T(z,t;\kappa)$ through the following Volterra equation of the second kind

$$aR_T+G^+*R_T=\partial_tG^--\frac{1}{4}\left(c_z+\frac{c\rho_z}{\rho}\right)G^+ \tag{42}$$

[in fact $R_T(z,t;\kappa)$ is the time derivative of the imbedding reflection kernel $R(z,t;\kappa)$], which has an unique solution for $R_T(z,t;\kappa)$ when the Green functions $G^{\pm}$ are known at the depth z. Using the PDEs for $G^{\pm}$ and the initial condition (27) for G^-, we can obtain the following useful condition (see Ref. 21 for a detailed derivation in a similar case)

$$\left[\left(\frac{\partial}{\partial t}R_T\right)(z,0;\kappa_1)-\left(\frac{\partial}{\partial t}R_T\right)(z,0;\kappa_2)\right]\Big/(\kappa_1^2-\kappa_2^2)=-\frac{1}{8}c^2\left(3c_z+\frac{c\rho_z}{\rho}\right). \tag{43}$$

Equations (27) (with $\kappa=\kappa_1$) and (43) can be used as two independent conditions for a simultaneous reconstruction [note that the right-hand sides of these two equations give two different combinations of c_z and ρ_z; c_z and ρ_z are reconstructed by the algorithm, while $c(z)$ and $\rho(z)$ are obtained by z integrations of c_z and ρ_z, respectively]. $(\frac{\partial}{\partial t}R_T)(z,0;\kappa)$ in Eq. (43) can be calculated numerically as follows:

$$\left(\frac{\partial}{\partial t}R_T\right)(z,0;\kappa)=\frac{R_T(z,\Delta t;\kappa)-R_T(z,0;\kappa)}{\Delta t},$$

where [cf. Eq. (42)]

$$R_T(z,0;\kappa)=\frac{1}{a}\left[\partial_tG^--\frac{1}{4}\left(c_z+\frac{c\rho_z}{\rho}\right)G^+\right](z,0;\kappa),$$

$$R_T(z,\Delta t;\kappa)=\frac{(\partial_tG^-)(z,\Delta t;\kappa)-\frac{1}{4}\left(c_z+\frac{c\rho_z}{\rho}\right)G^+(z,\Delta t;\kappa)-\frac{1}{2}\Delta tG^+(z,\Delta t;\kappa)R_T(z,0;\kappa)}{a+\frac{1}{2}\Delta tG^+(z,0;\kappa)}.$$

The reflection kernel $R(t;\kappa)$ for $\kappa=0.01$ and $\kappa=3$ are plotted in Fig. 6(a) for the profiles shown in Fig. 6(b). The simultaneous reconstruction of the velocity $c(z)$ and the density $\rho(z)$ using $R(t;\kappa=0.01)$ and $R(t;\kappa=3)$ are given by the circles (output with every fourth point for graphical clarity) in Fig. 6(b). The gridsize used here is $\Delta t=2\tau(z_{\max})/N$, $\Delta z=c(z)\Delta t/2$, with $N=512$, $z_{\max}=1.1$. Note that the simultaneous reconstruction algorithm works even for two small κ values, e.g., $\kappa=0.01$ and $\kappa=0.1$; however, the algorithm will then be less stable under noise. For this reason it is

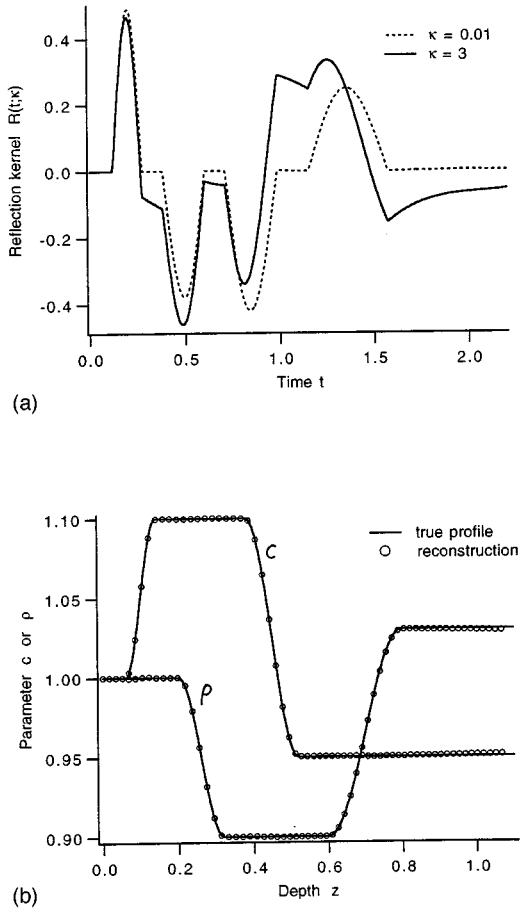


FIG. 6. (a) The reflection kernel $R(t;\kappa)$ for two different values of the Hankel-transform parameter κ, $\kappa_1=0.01$, and $\kappa_2=3$. (b) Simultaneous reconstruction of the velocity and density using the reflection data for $\kappa_1=0.01$, and $\kappa_2=3$.

better to choose the two values of κ to be separated well enough for the simultaneous reconstruction.

V. CONCLUSION

We have considered the acoustic inverse problem for a short-pulse point source over a stratified half-space. Transverse spatial Hankel transform has been applied to the transient reflected and incident pressures on the surface. The reflection kernel in the Hankel-transformed space has been obtained through a deconvolution. The time-domain Green function approach based on wave splitting has been used to reconstruct the parameters. Numerical results have been presented to illustrate the Hankel transform, the deconvolution, and the reconstruction for a synthetical example, in which the transient reflection data in the physical space is generated by the finite-difference time-domain (FDTD) method. The reconstruction algorithm has been shown numerically to be stable, and has been checked using a closed-form expression for the reflection kernel for a special density profile. Numerical results for a simultaneous reconstruction of the velocity and density using two values of the Hankel-transform parameter have also been given. Analogous acoustic reconstruction problems can be formulated for the case of oblique plane-wave incidence on a stratified half-space. This problem can also be solved by the wave-splitting Green function approach and the necessary formulas are given in an Appendix. Simultaneous reconstruction of the velocity and density then uses incidence for two different angles.

In most frequency-based approaches, one has to assume that the parameters become constant after a certain depth (since the reflectivity problem in the frequency domain has a solution only for such a profile); see, e.g., Refs. 26 and 16 (and the earlier references given there). An approach which treats the problem directly in the time domain (like the present approach) does not have such a restriction for the parameter profile. The main advantage of the present time-domain approach is that the reconstruction is theoretically exact and noniterative.

Parameter inversions from some experimental measurements using the present wave-splitting approach have been carried out recently at our department (Department of Electromagnetic Theory, Royal Institute of Technology, Sweden), cf. e.g., Refs. 27 and 28. However, the experimental aspects will not be addressed in the present paper.

ACKNOWLEDGMENTS

The authors are grateful to Gunnar Larson for providing his deconvolution code, and to Olle Brander and Staffan Ström for valuable discussions. The partial support of the Swedish Research Council for Engineering Sciences is also gratefully acknowledged.

APPENDIX: RECONSTRUCTION USING OBLIQUELY INCIDENT PLANE WAVES

In this Appendix, we consider the inverse acoustic problem using obliquely incident plane waves. If the plane of incidence is in the xz plane, then the pressure has the following form:

$$p=p\left(z,t-\frac{\sin\theta}{c_0}x;\theta\right), \tag{A1}$$

i.e.,

$$\partial_x=-\frac{\sin\theta}{c_0}\partial_t, \quad \partial_y=0, \tag{A2}$$

where θ is the incident angle, and c_0 is the velocity in the upper homogeneous half-space. Thus we only need to consider the pressures at those points with $x=0$, and the condition $x=0$ will be suppressed in all the equations hereafter. The pressures are assumed to be identically zero in the stratified half-space $z>0$ (with $x=0$) for the time $t\leqslant 0$, i.e.,

$$p(z,t;\theta)|_{t\leqslant 0}=0, \quad \text{for } z>0. \tag{A3}$$

Then the acoustic wave equation (1) becomes

$$\frac{1}{\overline{c^2}(z,\theta)}\frac{\partial^2 p}{\partial t^2}-\frac{\partial^2 p}{\partial z^2}+\frac{\rho_z}{\rho}\frac{\partial p}{\partial z}=0, \tag{A4}$$

where

$$\overline{c}(z;\theta)=1\Big/\sqrt{\frac{1}{c^2(z)}-\frac{\sin^2\theta}{c_0^2}}. \tag{A5}$$

The split pressures are as follows:

$$p^{\pm}(z,t;\theta)=\frac{1}{2}\left[p\mp\overline{c}\int_0^t\frac{\partial p}{\partial z}(z,t';\theta)dt'\right]. \tag{A6}$$

The following dynamic equation for the split pressures is then obtained

$$\frac{\partial}{\partial z}\begin{bmatrix}p^+\\p^-\end{bmatrix}=\begin{bmatrix}-(1/\overline{c})\dfrac{\partial}{\partial t}+\dfrac{\rho_z}{2\rho}+\dfrac{\overline{c_z}}{2\overline{c}} & -\dfrac{\rho_z}{2\rho}-\dfrac{\overline{c_z}}{2\overline{c}}\\ -\dfrac{\rho_z}{2\rho}-\dfrac{c_z}{2\overline{c}} & (1/\overline{c})\dfrac{\partial}{\partial t}+\dfrac{\rho_z}{2\rho}+\dfrac{\overline{c_z}}{2\overline{c}}\end{bmatrix}\times\begin{bmatrix}p^+\\p^-\end{bmatrix}, \tag{A7}$$

where

$$\overline{c_z}(z;\theta)\equiv\frac{\partial}{\partial z}\overline{c}=c_z(z)\cdot\frac{c_0^3}{[c_0^2-c^2(z)\sin^2\theta]^{3/2}}. \tag{A8}$$

Similarly, the Green functions are defined as follows:

$$\begin{bmatrix}p^+\\p^-\end{bmatrix}(z,t+\tau(z);\theta)=a(z)\begin{bmatrix}p^+(0,t;\theta)\\0\end{bmatrix}+\begin{bmatrix}G^+(z,t;\theta)*p^+(0,t;\theta)\\G^-(z,t;\theta)*p^+(0,t;\theta)\end{bmatrix}, \tag{A9}$$

where

$$\tau(z)=\int_0^z\frac{dz'}{\overline{c}(z')},\quad a(z)=\sqrt{\frac{\rho(z)\overline{c}(z)}{\rho(0)\overline{c}(0)}}. \tag{A10}$$

The PDEs for the Green functions are as follows:

$$\frac{\partial}{\partial z}G^+=\frac{1}{2}\left[\frac{\overline{c_z}}{\overline{c}}+\frac{\rho_z}{\rho}\right](G^+-G^-), \tag{A11}$$

$$\frac{\partial}{\partial z}G^-=\frac{2}{\overline{c}}\frac{\partial}{\partial t}G^--\frac{1}{2}\left[\frac{\overline{c_z}}{\overline{c}}+\frac{\rho_z}{\rho}\right](G^+-G^-). \tag{A12}$$

We also have the following initial and boundary conditions for the Green functions:

$$G^-(z,0;\theta)=\frac{1}{4}\,a\left\{\frac{c_0^3}{[c_0^2-c^2(z)\sin^2\theta]^{3/2}}\cdot c_z(z)+\frac{\overline{c\rho}_z}{\rho}\right\}, \tag{A13}$$

$$G^+(0,t;\theta)=0, \tag{A14}$$

$$G^-(0,t;\theta)=R(0,t;\theta). \tag{A15}$$

Equations (A11)–(A15) constitute a set of equations with which the inverse problem can be solved.

Example A1. This example shows simultaneous reconstruction using two different incident angles. The reflection kernel $R(t;\theta)$ for $\theta=0°$ and 45 ° are plotted in Fig. A1(a) for the profiles shown in Fig. A1(b) (where steep slopes are used to approximate jump discontinuities in the density). The simultaneous reconstruction of the velocity $c(z)$ and the density $\rho(z)$ are given by the dotted lines in Fig. A1(b), which essentially coincide with the true profiles (the solid lines). The grid size used here is $\Delta t=2\tau(z_{\max})/N$, $\Delta z=z_{\max}/N$, $\Delta t=2\Delta z/\overline{c}$, with $N=1024$, $z_{\max}=1.1$. Since the grid size Δt differs for different depth z, an interpolation is used in propagating $G^{\pm}$ to the next layer. We have also tested the noisy data for the reflection kernel $R(t;\theta)$ as shown by the dashed line in Fig. A1(a) [the Gaussian random noise has a root mean square signal to noise ratio (rms S/N) of 8.6]. The corresponding reconstruction is given by the dashed line in Fig. A1(b), which indicates that the inverse algorithm is stable for the noise data. One can certainly perform a similar reconstruction using the reflection kernel $R(t;\theta)$ for $\theta=0°$ and 30°. Again, we note that the incident angles need to be separated well enough for the simultaneous reconstruction to be stable under noise.

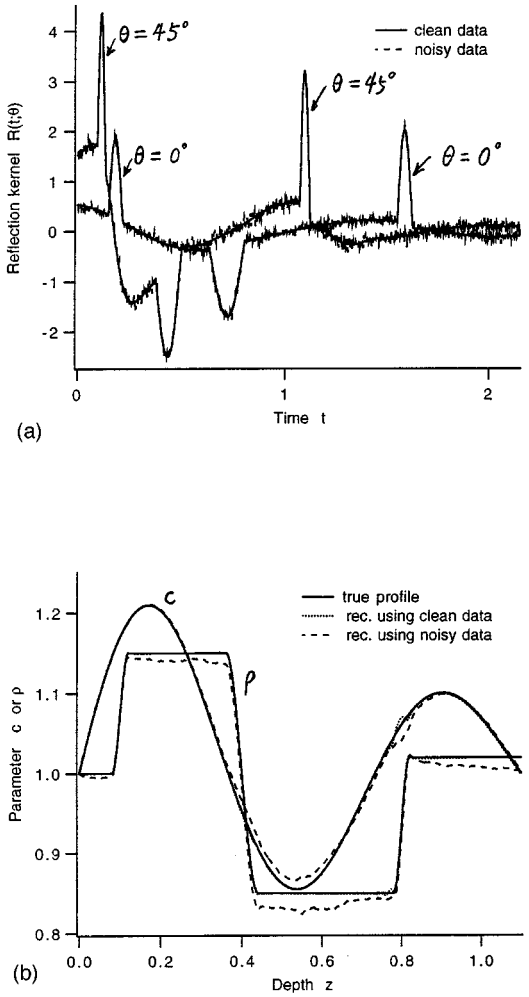


FIG. A1. (a) The reflection kernel $R(t;\theta)$ for two different incident angles $\theta=0°$ and 45°. (b) Simultaneous reconstruction of the velocity and density using the reflection data for $\theta=0°$ and 45°.

[1] J. P. Corones, M. E. Davison, and R. J. Krueger, ''Direct and inverse scattering in the time domain via invariant imbedding equations,'' J. Acoust. Soc. Am. **74**, 1535–1541 (1983).

[2] G. Kristensson and R. J. Krueger, ''Direct and inverse scattering in the time domain for a dissipative wave equation. Part I. Scattering operators,'' J. Math. Phys. (N.Y.) **27**, 1667–1682 (1986); ''Part II. Simultaneous reconstruction of dissipation and phase velocity profiles,'' J. Math. Phys. (N.Y.) **27**, 1683–1693 (1986); ''Part III. Scattering operators in the presence of a phase velocity mismatch,'' J. Math. Phys. (N.Y.) **28**, 360–70 (1987); ''Part IV. Use of phase velocity mismatches to simplify inversions,'' Inverse Probl. **5**, 375–388 (1989).

[3] V. H. Weston, ''Factorization of the wave equation in higher dimensions,'' J. Math. Phys. (N.Y.) **28**, 1061–1068 (1987).
[4] S. He and S. Ström, ''The electromagnetic scattering problem in the time domain for a dissipative slab and a point source using invariant imbedding,'' J. Math. Phys. (N.Y.) **32**, 3529–3539 (1991).
[5] S. He and V. H. Weston, ''Inverse problem for the dissipative wave equation in a stratified half-space and linearization of the imbedding equation,'' Inverse Probl. **8**, 435–455 (1992).
[6] R. J. Krueger and R. L. Ochs, ''A Green's function approach to the determination of internal fields,'' Wave Motion **11**, 525–543 (1989).
[7] A. Karlsson, H. Otterheim, and R. Stewart, ''Electromagnetic fields in an inhomogeneous plasma from obliquely incident transient plane waves,'' Radio Sci. **28**, 365–378 (1993).
[8] L. Fishman, ''Exact and operator rational approximation solutions of the Helmholtz, Weyl composition equation in underwater acoustic—The quadratic profile,'' J. Math. Phys. (N.Y.) **33**, 1887–1914 (1992).
[9] S. He, ''A time-harmonic Green function technique and wave propagation in a stratified nonreciprocal chiral slab with multiple discontinuities,'' J. Math. Phys. (N.Y.) **33**, 4103–4110 (1992).
[10] S. He and Y. Hu, ''Electromagnetic scattering for a stratified biisotropic (nonreciprocal chiral) slab: Numerical computations,'' IEEE Trans. Antennas Propag. **41**, 1057–1062 (1993).
[11] S. He, S. I. Kabanikhin, V. G. Romanov, and S. Ström, ''Analysis of the Green function approach to one-dimensional inverse problems,'' J. Math. Phys. (N.Y.) **34**, 5724–5746 (1993).
[12] J. Cao and S. He, ''Closed-form solution for the transient reflected pressure for a point source above an acoustic half-space with an exponentially stratified density,'' J. Acoust. Soc. Am. **96**, 2516–2525 (1994).
[13] D. W. Oldenberg, T. Scheuer, and S. Levy, ''Recovery of the acoustic impedance from reflection seismograms,'' Geophysics **48**, 1318–1337 (1983).
[14] W. W. Symes, ''Impedance profile inversion via the first transport equation,'' J. Math. Anal. Appl. **94**, 435–453 (1983).
[15] A. E. Yagle and B. Levy, ''A fast algorithm solution of the inverse problem for a layered acoustic medium probed by spherical harmonic waves,'' J. Acoust. Soc. Am. **78**, 729–737 (1985).
[16] L. Amundsen and B. Ursin, ''Frequency-wave number inversion of acoustic data,'' Geophysics **56**, 1027–1039 (1991).
[17] H. L. Deng, ''Acoustic-wave propagation in thin-layered media with steep reflectors,'' Geophysics **59**, 1593–1604 (1994).
[18] V. H. Weston, ''Invariant imbedding and wave splitting in R^3. Part II. the Green function approach to inverse scattering,'' Inverse Probl. **8**, 919–947 (1992).
[19] V. H. Weston and S. He, ''Wave-splitting of the telegraph equation in R^3 and its application to the inverse scattering,'' Inverse Probl. **9**, 789–812 (1993).
[20] F. G. Friedlander, *Sound Pulses* (Cambridge U.P., Cambridge, 1958).
[21] S. He and A. Karlsson, ''Time domain Green functions technique for a point source over a dissipative stratified half-space,'' Radio Sci. **28**, 513–526 (1993).
[22] S. He, Y. Hu, and S. Ström, ''Time domain Green functions technique for a point source over a dissipative stratified half-space with a phase velocity mismatch at the surface,'' Wave Motion **17**, 241–254 (1993).
[23] K. S. Yee, ''Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media,'' IEEE Trans. Antennas Propag. **14**, 302–307 (1966).
[24] R. M. Alford, K. R. Kelly, and D. M. Boore, ''Accuracy of finite-difference modeling of the acoustic wave equation,'' Geophysics **39**, 834–842 (1974).
[25] N. S. Nahman and M. E. Guillaume, ''Deconvolution of time domain waveforms in the presence of noise,'' NBS Technical Note 1047, October 1981.
[26] S. S. Pan, R. A. Phinney, and R. I. Odem, ''Full-waveform inversion of plane-wave seismograms in stratified acoustic media: Theory and feasibility,'' Geophysics **53**, 21–31 (1988).
[27] F. Axelsson and P. Rönnqvist, ''Investigation on the possibility to reconstruct an acoustic impedance profile using ultrasonic sound,'' Technical Report TRITATET-EX 93-5, Department of Electromagnetic Theory, Royal Institute of Technology, Stockholm, Sweden, 1990.
[28] P. Fuks, A. Karlsson, and G. Larson, ''Direct and inverse scattering from dispersive media,'' Inverse Probl. **10**, 555–571 (1994).

Extraction of a target scattering response from measurements made over long ranges in shallow water

Angie Sarkissian
Naval Research Laboratory, Washington, DC 20375-5350

(Received 16 October 1996; accepted for publication 24 December 1996)

The scattering response of a structure may be extracted from acoustic measurements made in shallow water at relatively long ranges from the target. The procedure described uses an array of sources to approximately generate a low-order mode and an array of receivers to isolate from the received echo a low-order mode. Since low-order modes behave as plane waves in the vicinity of the target, the free-field target response may be extracted from such measurements. In general, such a procedure for the generation of a low-order mode requires a large number of sources. The number is of the order of the number of propagating, nonevanescent modes present in the shallow water environment at the frequencies used. If this number is prohibitively large, temporal discrimination may be used in addition to beamforming to reduce the number of sources necessary for the isolation of the free-field target response. Similarly, the number of receivers required for the isolation of a mode from the received signal may be reduced by the use of temporal discrimination in addition to beamforming techniques. This article presents the detailed procedure used as well as numerical simulations to demonstrate the feasibility of such a process. [S0001-4966(97)00208-7]

PACS numbers: 43.30.Bp, 43.30.Gv, 43.30.Vh [JHM]

INTRODUCTION

A single omnidirectional source placed in shallow water will produce an incident field on a target that contains interfering multipath arrivals. Similarly, the scattered field from the target arrives at a receiver through multipaths which produce interference effects. If the source and the receiver are placed at long ranges from the target, the multipaths arrive at the receiver very close in time. Thus the use of simple time-gating techniques will not be feasible for isolation of the target response from the received signal.

The use of beamforming techniques allows the isolation of the target response more readily.[1] In this case, beamforming may be applied to a vertical or horizontal array of sources to generate a low-order shallow water mode since low-order modes, in the vicinity of the target, behave as plane waves. Similarly, a vertical or a horizontal array of receivers may be used to extract from the received signal a low-order mode. The procedure for the generation of a given mode in shallow water with a discrete vertical or horizontal array of sources is well known but it typically requires the number of sources to be at least as many as the number of propagating modes present in the shallow water medium at the frequencies of interest. This number, which increases with frequency and with water depth, may be prohibitively large. The alternative approach used here is to generate a low-order mode and suppress other low-order modes while the higher-order modes are removed by the use of time-gating techniques.

I. THE SCATTERED FIELD IN A BOUNDED MEDIUM

In shallow water, the acoustic pressure field $P(\mathbf{r},t)$ satisfies the wave equation

$$\rho\nabla\cdot\left(\frac{1}{\rho}\nabla P\right)-\frac{1}{c^2}\frac{\partial^2 P}{\partial t^2}=0, \tag{1}$$

where $c=c(\mathbf{r})$ is the time-independent sound speed. The use of Fourier decomposition reduces this to a three-dimensional problem,

$$P(\mathbf{r},t)=\int p(\mathbf{r},\omega)e^{-i\omega t}\,d\omega, \tag{2}$$

where the solution in the frequency domain $p(\mathbf{r},\omega)$ satisfies

$$\rho\nabla\cdot\left(\frac{1}{\rho}\nabla p\right)+k^2 p=0. \tag{3}$$

Above, $k=\omega/c$ is the wave number.

An omnidirectional source placed at position $\mathbf{r}_s$ produces a field that satisfies

$$\rho\nabla\cdot\left(\frac{1}{\rho}\nabla p\right)+k^2 p=q_0\delta(\mathbf{r}-\mathbf{r}_s), \tag{4}$$

where q_0 is the source amplitude and δ is the Dirac delta function. In a range-independent environment, where the coordinate system is placed with the z axis along the vertical direction with $c=c(z)$, the normal mode solution to the above equation is[2]

$$p(\mathbf{r})=q_0\pi i\sum_{n=1}^{\infty}u_n(z)u_n(z_s)H_0^{(1)}(k_n|\boldsymbol{\rho}-\boldsymbol{\rho}_s|), \tag{5}$$

where $H_0^{(1)}$ is the zeroth-order cylindrical Hankel function of the first kind and $\boldsymbol{\rho}$ is the projection of $\mathbf{r}$ onto the x-y plane. At long ranges, where $k|\boldsymbol{\rho}-\boldsymbol{\rho}_s|\gg 1$, only a finite number N of propagating modes contribute to the above sum. They form the discrete spectrum of the nonevanescent modes with

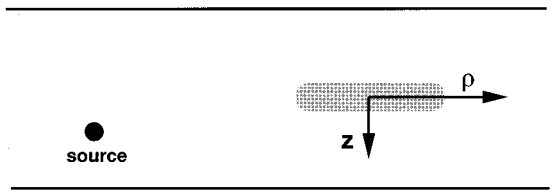


FIG. 1. Geometry showing the source, the scatterer, and the coordinate system in the shallow water environment.

real k_n values. In this case the far-field approximation to the Hankel function gives

$$p(\mathbf{r}) = q_0(2\pi i)^{1/2}\sum_{n=1}^{N} \frac{1}{(k_n|\boldsymbol{\rho}-\boldsymbol{\rho}_s|)^{1/2}} \times u_n(z)u_n(z_s)e^{i(k_n|\boldsymbol{\rho}-\boldsymbol{\rho}_s|)}. \tag{6}$$

Consider the geometry shown in Fig. 1, where an omnidirectional source produces a field satisfying Eq. (4), which is then scattered by a target and measured by a receiver. If both the source and the receiver are distant from the target, and the target is sufficiently distant from shallow water boundaries for multiple scattering effects to be negligible,[3,4] then the scattered field may be written in a simple form.[5,6] Ingenito[5] derived this algorithm for an arbitrary target satisfying rigid boundary conditions and stated that it may be generalized to the more generic case. Here the algorithm is derived for a target satisfying arbitrary boundary conditions. We follow a procedure similar to Ingenito's.

Assuming the sound speed near the scatterer is approximately constant, each normal mode u_n in this region varies approximately sinusoidally,

$$u_n(z) = a_n e^{i\alpha_n z} + b_n e^{-i\alpha_n z}, \tag{7}$$

where

$$\alpha_n = \left(\frac{\omega^2}{c^2} - k_n^2\right)^{1/2}, \tag{8}$$

$$a_n = \frac{1}{2}\left[u_n(z_c) + \frac{u_n'(z_c)}{i\alpha_n}\right]e^{-i\alpha_n z_c}, \tag{9}$$

$$b_n = \frac{1}{2}\left[u_n(z_c) - \frac{u_n'(z_c)}{i\alpha_n}\right]e^{i\alpha_n z_c}, \tag{10}$$

z_c is the z coordinate of the center of the scatterer, and prime designates differentiation with respect to the argument. For simplicity we choose the center of the scatterer to lie at the coordinate origin, with $z_c = 0$, without any loss of generality. Substitution of Eq. (7) into the expression for the incident field p_i given by Eq. (6) results in

$$p_i(\mathbf{r}) = q(2\pi i)^{1/2}\sum_{n=1}^{N} \frac{u_n(z_s)}{(k_n|\boldsymbol{\rho}-\boldsymbol{\rho}_s|)^{1/2}}\left[a_n e^{i(k_n|\boldsymbol{\rho}-\boldsymbol{\rho}_s|+\alpha_n z)} + b_n e^{i(k_n|\boldsymbol{\rho}-\boldsymbol{\rho}_s|-\alpha_n z)}\right]. \tag{11}$$

This represents the expansion of the incident field in terms of a finite number of plane waves incident from various directions. The shallow water Green's function, $G(\mathbf{r},\mathbf{r}')$, for $k|\mathbf{r}-\mathbf{r}'| \gg 1$, which is also the solution to Eq. (4) with $q_0 = 1$, may similarly be expanded in terms of plane waves,

$$G(\mathbf{r},\mathbf{r}') = (2\pi i)^{1/2}\sum_{n=1}^{N} \frac{u_n(z')}{(k_n|\boldsymbol{\rho}-\boldsymbol{\rho}'|)^{1/2}} \times\left[a_n e^{i(k_n|\boldsymbol{\rho}-\boldsymbol{\rho}'|+\alpha_n z)} + b_n e^{i(k_n|\boldsymbol{\rho}-\boldsymbol{\rho}'|-\alpha_n z)}\right]. \tag{12}$$

To solve for the scattered field p_s, we use the Helmholtz integral equation[7]

$$p_s(\mathbf{r}') = \frac{1}{4\pi}\int_S [p(\mathbf{r})\hat{n}\cdot\boldsymbol{\nabla}G(\mathbf{r},\mathbf{r}') - \hat{n}\cdot\boldsymbol{\nabla}p(\mathbf{r})G(\mathbf{r},\mathbf{r}')]dS, \tag{13}$$

where $p(\mathbf{r})$ is the total field,

$$p(\mathbf{r}) = p_i(\mathbf{r}) + p_s(\mathbf{r}). \tag{14}$$

We next define the variables $p^b_{n\pm}$ to be the the scattered field produced by a plane wave of the form $e^{i(k_n|\boldsymbol{\rho}-\boldsymbol{\rho}_s|\pm\alpha_n z)}$ incident on the scatterer in the bounded medium. Using the principle of the wave superposition for a linear medium, the incident field given by Eq. (11) produces a scattered field given by

$$p_s(\mathbf{r}) = q(2\pi i)^{1/2}\sum_{n=1}^{N} \frac{u_n(z_s)}{(k_n\rho_s)^{1/2}}\left[a_n p^b_{n+} + b_n p^b_{n-}\right]. \tag{15}$$

We have assumed above that the size of the scatterer is much less than the distance between the source and the scatterer, so that $|\boldsymbol{\rho}-\boldsymbol{\rho}_s|$ remains constant for points $\boldsymbol{\rho}$ on the region near the surface of the scatterer.

An expression for $p^b_{n\pm}$ may be determined from the Helmholtz integral equation. The substitution of the Green's function expression, Eq. (12), into Eq. (13) and the use of Eq. (14) results in

$$\begin{aligned} p^b_{n\pm}(\mathbf{r}') = \frac{1}{2}\left(\frac{i}{2\pi}\right)^{1/2}\sum_{m=1}^{N} \frac{u_m(z')}{(k_m\rho')^{1/2}} \times\int_S \{&[e^{i(k_n|\boldsymbol{\rho}-\boldsymbol{\rho}_s|\pm\alpha_n z)} + p^b_{n\pm}]\hat{n}\cdot\boldsymbol{\nabla}[a_m e^{i(k_m|\boldsymbol{\rho}-\boldsymbol{\rho}'|+\alpha_m z)} \\ &+ b_m e^{i(k_m|\boldsymbol{\rho}-\boldsymbol{\rho}'|-\alpha_m z)}] - \hat{n}\cdot\boldsymbol{\nabla}[e^{i(k_n|\boldsymbol{\rho}-\boldsymbol{\rho}_s|\pm\alpha_n z)} + p^b_{n\pm}][a_m e^{i(k_m|\boldsymbol{\rho}-\boldsymbol{\rho}'|+\alpha_m z)} \\ &+ b_m e^{i(k_m|\boldsymbol{\rho}-\boldsymbol{\rho}'|-\alpha_m z)}]\}dS, \end{aligned} \tag{16}$$

where we have again assumed that the size of the scatterer is much less than the distance between $\mathbf{r}'$ and the scatterer, so that $|\boldsymbol{\rho}-\boldsymbol{\rho}'|$ remains constant for points $\boldsymbol{\rho}$ on the surface of the scatterer.

We define the variables $p^u_{n\pm}$ to be the scattered field produced by a plane wave field of the form $e^{i(k_n|\boldsymbol{\rho}-\boldsymbol{\rho}_s|\pm\alpha_n z)}$ incident on the target in free space. This field also satisfies the Helmholtz integral equation but with the free-field Green's function,

$$p^u_{n\pm}(\mathbf{r}')=\frac{1}{4\pi}\int_S\Big\{[e^{i(k_n|\boldsymbol{\rho}-\boldsymbol{\rho}_s|\pm\alpha_n z)}+p^u_{n\pm}(\mathbf{r})]\hat{n}\cdot\nabla\frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}-\hat{n}\cdot\nabla[e^{i(k_n|\boldsymbol{\rho}-\boldsymbol{\rho}_s|\pm\alpha_n z)}+p^u_{n\pm}(\mathbf{r})]\frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}\Big\}dS. \tag{17}$$

This has the far-field approximation given by

$$p^u_{n\pm}(\mathbf{r}')=\frac{1}{4\pi}\frac{e^{ikr'}}{r'}\int_S\{[e^{i(k_n|\boldsymbol{\rho}-\boldsymbol{\rho}_s|\pm\alpha_n z)}+p^u_{n\pm}(\mathbf{r})]\hat{n}\cdot\nabla e^{-ik\hat{r}'\cdot\mathbf{r}}-\hat{n}\cdot\nabla[e^{i(k_n|\boldsymbol{\rho}-\boldsymbol{\rho}_s|\pm\alpha_n z)}+p^u_{n\pm}(\mathbf{r})]e^{-ik\hat{r}'\cdot\mathbf{r}}\}dS, \tag{18}$$

or

$$p^u_{n\pm}(r,\theta,\phi)=\frac{e^{ikr}}{r}e^{i(k_n\rho_s)}S(\hat{r},-\hat{k}_{n\pm}), \tag{19}$$

where $(\hat{k}_{n\pm})$ is the direction of the plane-wave $e^{i(k_n|\boldsymbol{\rho}-\boldsymbol{\rho}_s|\pm\alpha_n z)}$ satisfying

$$\frac{\omega}{c}\hat{k}_{n\pm}\cdot\mathbf{r}=(k_n|\boldsymbol{\rho}-\boldsymbol{\rho}_s|\pm\alpha_n z); \tag{20}$$

S is the scattering function in an unbounded medium that depends only on the direction of the incident field and the direction of the position where the scattered field is measured. Combining Eqs. (16), (18), and (19), and assuming that the scattered field on the surface of the target and its normal derivative in a bounded medium are the same as the fields in an unbounded medium, i.e., $p^u_{n\pm}|_S=p^b_{n\pm}|_S$ and $\hat{n}\cdot\nabla p^u_{n\pm}|_S=\hat{n}\cdot\nabla p^b_{n\pm}|_S$, we obtain

$$p^b_{n\pm}(\mathbf{r})=(2\pi i)^{1/2}\sum_{m=1}^{N}\frac{u_m(z)}{(k_m\rho)^{1/2}}e^{i(k_n\rho_s+k_m\rho)}\times[a_m S(\hat{k}_{m+},\hat{k}_{n\pm})+b_m S(\hat{k}_{m-},\hat{k}_{n\pm})]. \tag{21}$$

Substituting this into Eq. (15) gives the total scattered field in the bounded medium

$$p_s(\mathbf{r})=2\pi i q\sum_{n=1}^{N}\sum_{m=1}^{N}\frac{u_n(z_s)u_m(z)}{(k_n\rho_s)^{1/2}(k_m\rho)^{1/2}}e^{i(k_n\rho_s+k_m\rho)}\times[a_n a_m S(\hat{k}_{m+},\hat{k}_{n+})+a_n b_m S(\hat{k}_{m-},\hat{k}_{n+})+b_n a_m S(\hat{k}_{m+},\hat{k}_{n-})+b_n b_m S(\hat{k}_{m-},\hat{k}_{n-})], \tag{22}$$

where a_n and b_n are given by Eqs. (9) and (10) with $z_c=0$.

If multiple scattering effects cannot be neglected because the target is close to a shallow water boundary, then other approaches for the determination of the scattered field include the T-matrix method,[8–10] the boundary-element method,[4,11] and the method of wave superposition.[12]

II. BEAMFORMING

The low-order modes in Eq. (5) have small α_n values. From Eqs. (7) and (8), it is apparent that, for these modes, the wave number in the horizontal direction k_n approaches ω/c, making the modes behave similar to plane waves propagating horizontally. The isolation of a target response could therefore be accomplished by the generation of one low-order mode and the extraction from the received echo of one low-order mode. The generation of a low-order mode may be accomplished with the use of a vertical or a horizontal array with source strengths weighed to suppress other low-order modes. Additionally, since higher-order modes have much lower group velocities in the horizontal direction, the use of time-windowing techniques may be used to isolate the low-order mode from the high-order modes. The isolation of a low-order mode from the received signal is similarly accomplished with an array of receivers and the use of time-windowing techniques.

A. Vertical arrays

Consider a vertical array of M sources with equal spacing Δz_s between any two adjacent sources, where source 1 has coordinates (ρ_s, z_s, ϕ_s). At long ranges, such a source distribution will produce a field given by

$$p(\mathbf{r})=(2\pi i)^{1/2}\sum_{m=1}^{M}q_m\sum_{n=1}^{N}\frac{1}{(k_n|\boldsymbol{\rho}-\boldsymbol{\rho}_s|)^{1/2}}u_n(z)\times u_n(z_s+[m-1]\Delta z_s)e^{i(k_n|\boldsymbol{\rho}-\boldsymbol{\rho}_s|)}, \tag{23}$$

where q_m is the source strength of source m, and we recall that N is the number of propagating modes in the environment.

If $M\geqslant N$, then such an array may be used to generate a field consisting of only one mode n, having the form

$$p_{\text{ideal}}(\mathbf{r})=Q(2\pi i)^{1/2}\frac{1}{(k_n|\boldsymbol{\rho}-\boldsymbol{\rho}_s|)^{1/2}}u_n(z)e^{i(k_n|\boldsymbol{\rho}-\boldsymbol{\rho}_s|)}. \tag{24}$$

The satisfaction of Eqs. (23) and (24) for all z requires the solution to the system of equations

$$\sum_{m=1}^{M}A_{nm}q_m=B_n, \tag{25}$$

where

$$A_{nm}=(2\pi i)^{1/2}\frac{1}{(k_n|\boldsymbol{\rho}-\boldsymbol{\rho}_s|)^{1/2}}\times u_n(z_s+[m-1]\Delta z_s)e^{i(k_n|\boldsymbol{\rho}-\boldsymbol{\rho}_s|)} \tag{26}$$

and

$$B_n=(2\pi i)^{1/2}Q\frac{1}{(k_n|\boldsymbol{\rho}-\boldsymbol{\rho}_s|)^{1/2}}e^{i(k_n|\boldsymbol{\rho}-\boldsymbol{\rho}_s|)}. \tag{27}$$

Matrix $\mathbf{A}$, with elements A_{nm}, may be inverted using the least-squares method to obtain a solution for the source strengths q_m,

$$\mathbf{q}=(\mathbf{A}^\dagger\mathbf{A})^{-1}\mathbf{A}^\dagger\mathbf{B}, \tag{28}$$

FIG. 2. The cylindrical scatterer where the dashed line on the right represents the plane of symmetry of the cylinder.

where † designates the adjoint operation; column vector **q** is defined to have elements q_n; and column vector **B** has element B_n in the nth row, and zeroes elsewhere. This procedure is similar to the determination of mode strengths from measurements made using an array of M hydrophones discussed by Tindel *et al.*[13]

If $M<N$, then the system of equations given by Eq. (25) will be overdetermined. In this case, the least-squares procedure described by Eq. (28) may still be used to generate a field with least-square difference from the ideal field described by Eq. (24). The alternative is to suppress $M-1$ low-order modes. This is accomplished by defining an $M\times M$ matrix $\mathbf{A}'$ with elements $A'_{nm}=A_{nm}$ and inverting this smaller matrix instead,

$$\mathbf{q}=(\mathbf{A}'^{\dagger}\mathbf{A}')^{-1}\mathbf{A}'^{\dagger}\mathbf{B}. \tag{29}$$

This procedure will place no restrictions on the higher-order modes $n>M$ which may instead be removed by temporal discrimination.

Using the principle of reciprocity, it is easy to show that beamforming with the receiver array to isolate a low-order mode while suppressing others requires the identical weighting on the receivers as the source strengths required for the generation of the mode described above.

The procedure is applied to the ribbed cylindrical nickel shell with hemispherical endcaps, shown in Fig. 2, having radius $a=3$ m. The cylinder has an axis of symmetry and a plane of symmetry, shown by the dashed line in Fig. 2. It has a total length $L=14a$, a shell thickness of $t=0.0074a$, and contains 85 ribs with rib spacing $d=0.14a$. The shorter ribs have a length of $l_1=0.077a$ and a thickness $t_1=0.0065a$. Every thirteenth rib is longer, having a length of $l_2=0.14a$ and a thickness of $t_2=0.0074a$. The free-field target response is simulated numerically using the program SARA-2D,[14] which models the structure using the finite elements method, and models the unbounded medium exterior to the structure using the finite and infinite elements method. For additional details on the simulation of the free-field target response, the reader is referred to Refs. 15 and 16.

The cylinder is placed in the shallow water environment, shown in Fig. 3 at a depth of 45 m. The normal modes in the shallow water are computed using the program KRAKEN.[17] Equation (22) is then applied to determine the response of the cylinder in this environment.

A vertical array of ten sources, at a distance of 10 km from the cylinder, is designed to generate mode 1 while suppressing modes 2–10, as described above with a value of $Q=1$ in Eq. (27). A ten receiver array, placed at the same location as the source array, measures the field scattered by the target. The receiver array is beamformed to extract from the received signal mode 1, while removing modes 2–10.

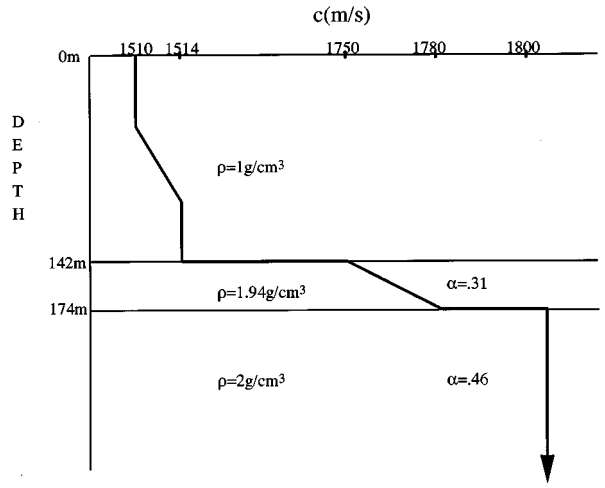


FIG. 3. The shallow water environment where α is the attenuation in dB per wavelength.

Figure 4 shows this beamformed signal in the frequency domain. The horizontal axis represents the cylinder orientation. It is the angle between the direction of the incident field and the axis of symmetry of the cylinder. It ranges from 0°, which represents end incidence, to 90°, which represents normal incidence. The vertical axis in Fig. 4 represents the

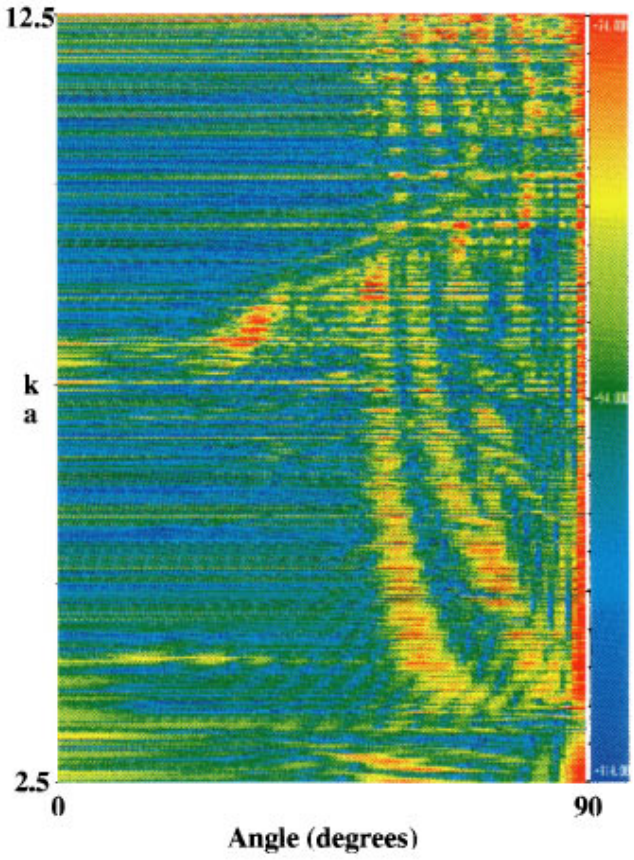


FIG. 4. Frequency domain response of the cylinder in shallow water with vertical array of sources and vertical array of receivers.

dimensionless quantity ka. Figure 5(a) shows the same field in the time domain. It is compared to the monostatic time domain response of the same cylinder in free space, shown in 5(b). Time in both plots is shown using the dimensionless quantity ct/a, where c is the speed of sound in the fluid region near the scatterer. The time origin in both curves is arbitrarily chosen to occur at some time prior to the arrival of the target response. The amplitude of both signals is normalized to have a maximum value of unity or 0 dB and plotted such that all values between -20 and 0 dB are shown in red.

Comparison of Fig. 5(a) to Fig. 5(b) shows that during the early time interval of $100a/c$, the shallow water response, Fig. 5(a), looks similar to the free-field response, Fig. 5(b). This portion of the field is generated by mode 1 incident on the target, where it behaves as a plane wave, while the receiver isolates mode 1 from the received signal which also behaves as a plane wave near the target. We will call this the mode 1–mode 1 arrival. Later the higher-order modes begin to arrive. They are modes 11 and higher that arrive with a time delay from the mode 1–mode 1 arrival. These higher-order modes are next removed using a time window of $31a/c$ and the windowed signal is inverted to the frequency domain. The result is shown in Fig. 6(a). Figure 6(b) shows the result of the same operation applied to the free-field response. We observe both plots are very similar in shape.

B. Horizontal arrays

In shallow water, a given mode may be generated by the use of a horizontal array instead, but the requirements in this case are different from those of a vertical array mainly because the spread in wave numbers k_n in the horizontal direction is different from the spread of wave numbers α_n in the vertical direction. While the wave numbers in the vertical direction α_n vary from low values to nearly the acoustic wave number $k=\omega/c$, the wave numbers in the horizontal direction k_n remain nearly equal to k. This makes beamforming more difficult with horizontal arrays and necessitates large spacing between adjacent sources or adjacent receivers for this case.

Consider a horizontal array of M sources where source m has coordinates $\mathbf{r}_m=(\rho_m,z_s,\phi_s)$. At large distances from the array, the field generated by this array is given by

$$p(\mathbf{r})=(2\pi i)^{1/2}\sum_{m=1}^{M} q_m \sum_{n=1}^{N} \frac{1}{(k_n|\boldsymbol{\rho}_m-\boldsymbol{\rho}_s|)^{1/2}} \times u_n(z)u_n(z_s)e^{i(k_n|\boldsymbol{\rho}_m-\boldsymbol{\rho}_s|)}, \tag{30}$$

where q_m is the source strength of mode m, and $\boldsymbol{\rho}_m$ is the projection of $\mathbf{r}_m$ onto the x-y plane.

The procedure for the generation of a mode with a horizontal array is similar to that with a vertical array. In this case, the ideal mode n to be generated has the form

$$p_{\text{ideal}}(\mathbf{r})=Q(2\pi i)^{1/2}\frac{1}{(k_n|\boldsymbol{\rho}-\boldsymbol{\rho}_1|)^{1/2}}u_n(z)e^{i(k_n|\boldsymbol{\rho}-\boldsymbol{\rho}_1|)}. \tag{31}$$

The satisfaction of Eqs. (30) and (31) for all z requires the solution of the system of equations

$$\sum_{m=1}^{M} C_{nm}q_m=D_n, \tag{32}$$

where

$$C_{nm}=(2\pi i)^{1/2}\frac{1}{(k_n|\boldsymbol{\rho}_m-\boldsymbol{\rho}_s|)^{1/2}}u_n(z_s)e^{i(k_n|\boldsymbol{\rho}_m-\boldsymbol{\rho}_s|)}, \tag{33}$$

and

$$D_n=(2\pi i)^{1/2}Q\frac{1}{(k_n|\boldsymbol{\rho}-\boldsymbol{\rho}_1|)^{1/2}}e^{i(k_n|\boldsymbol{\rho}-\boldsymbol{\rho}_1|)}. \tag{34}$$

Again, for $M\geqslant N$, matrix $\mathbf{C}$ with elements C_{nm} may be inverted using the least-squares method to obtain a solution for the source strengths.

$$\mathbf{q}=(\mathbf{C}^{\dagger}\mathbf{C})^{-1}\mathbf{C}^{\dagger}\mathbf{D}, \tag{35}$$

where column vector $\mathbf{D}$ has element D_n in the nth row and zeroes in the remaining rows. For $M<N$, in order to suppress $M-1$ low-order modes, an $M\times M$ matrix C', defined to have elements $C'_{nm}=C_{nm}$, is inverted instead,

$$\mathbf{q}=(\mathbf{C}'^{\dagger}\mathbf{C}')^{-1}\mathbf{C}'^{\dagger}\mathbf{D}. \tag{36}$$

Numerical simulations are performed for the horizontal array with the same scatterer and same environment used for the vertical array simulations. The horizontal array contains 21 sources with equal spacing of 400 m between adjacent sources. A vertical array of ten receivers is placed at the same location as the source that is farthest from the scatterer where the distance between this receiver array and the scatterer is 10 km.

Figure 7(a) and (b) shows the time domain impulse response in shallow water and free space, respectively, both in the frequency range corresponding to $2.5\leqslant ka\leqslant 2.9$. Again, the time origin is arbitrary and both plots are normalized to have a maximum value of 0 dB and plotted such that all values between -20 and 0 dB are shown in red. In shallow water, during the early period, we observe that the mode 1–mode 1 arrival is very similar to the free-field response. At later times, modes 22 and higher produce additional effects in Fig. 7(a). The use of a time window of width $155a/c$ followed by a Fourier transform applied to Fig. 7(a) and (b) produces the frequency domain response shown in Fig. 8(a) and 8(b), respectively. Again, the shallow water response, shown in Fig. 8(a), is very similar to the free-field response, shown in Fig. 8(b).

Using the principle of reciprocity, we note that the above simulation is identical to the case where the sources and the receivers are interchanged, i.e., the case where a vertical array of ten sources produces a field that is scattered by the cylinder and measured by a horizontal array of 21 receivers.

III. CONCLUSION

The scattering response of a target may be extracted from measurements made in shallow water using modal decomposition with vertical or horizontal arrays of sources and receivers. If the number of sources or receivers are less than the number of propagating modes present in the shallow wa-

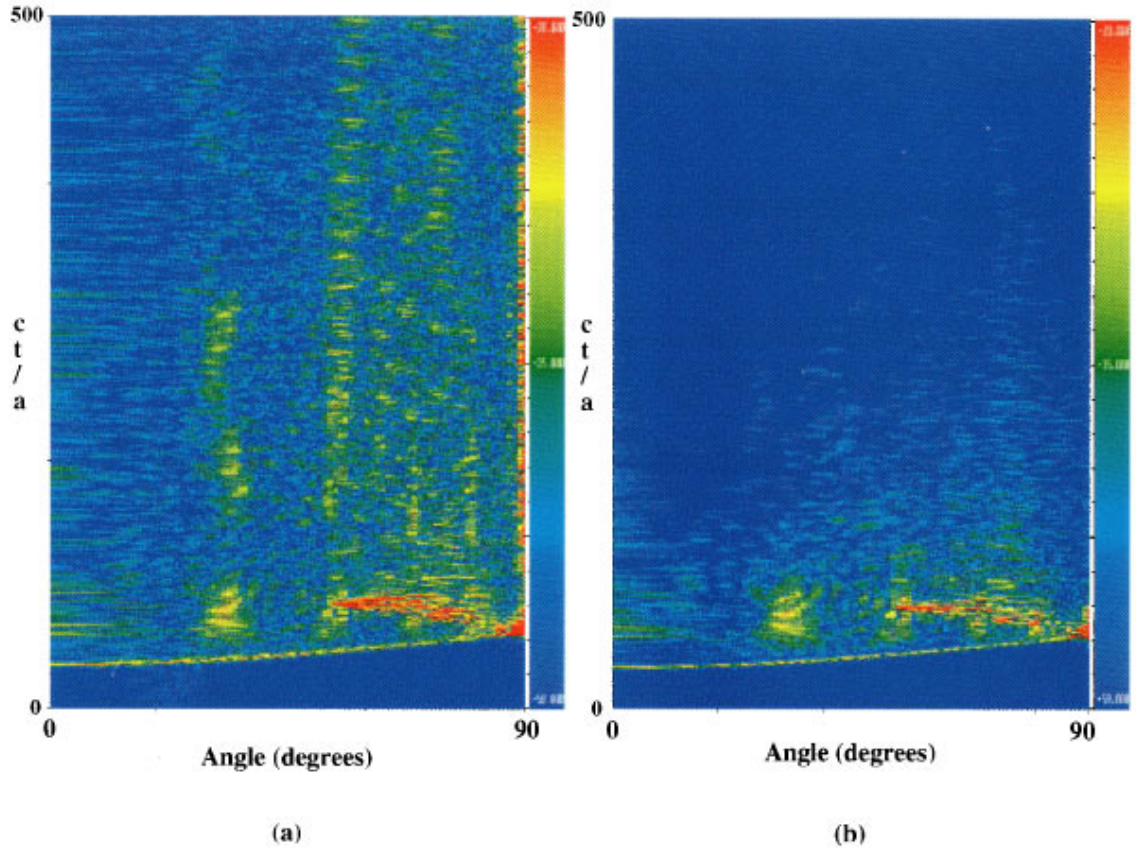


FIG. 5. Time domain response of cylinder for $2.5 \leqslant ka \leqslant 12.5$ in (a) shallow water with vertical array of sources and vertical array of receivers and (b) free space.

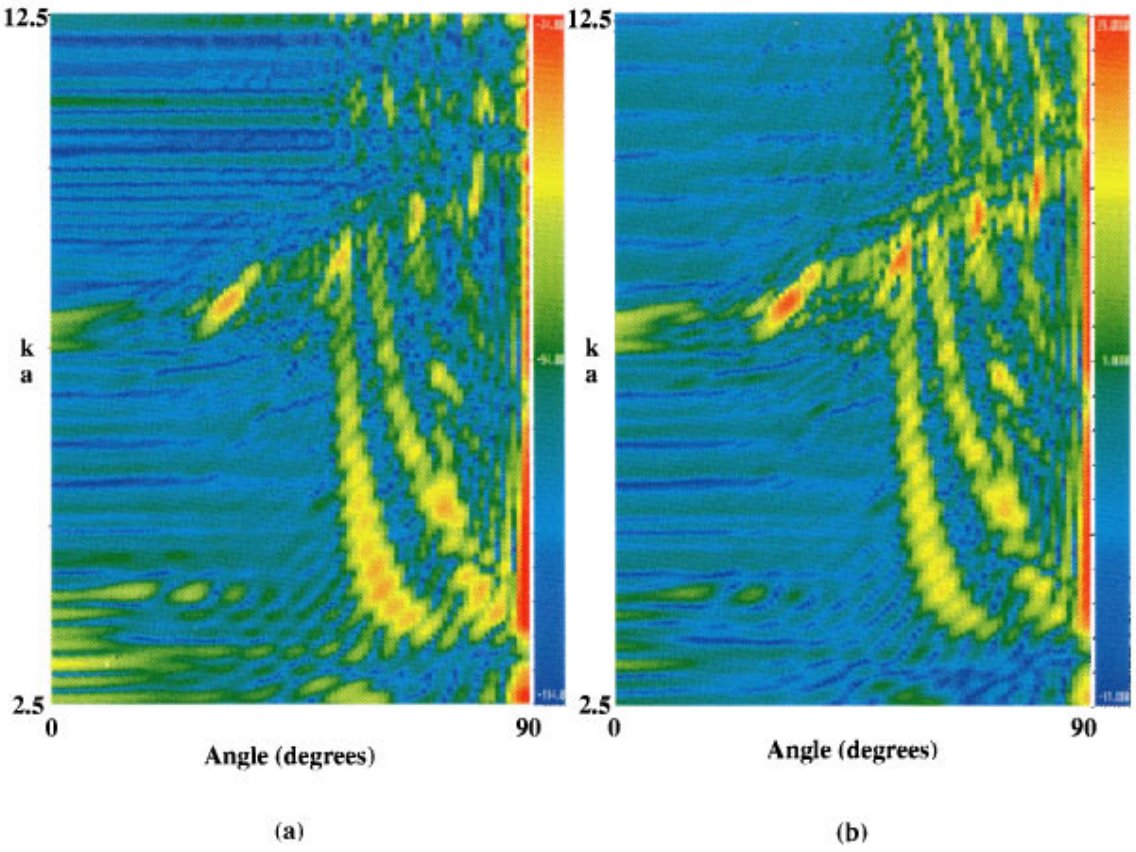


FIG. 6. Frequency domain response of cylinder in (a) shallow water with vertical array of sources and vertical array of receivers and (b) free space.

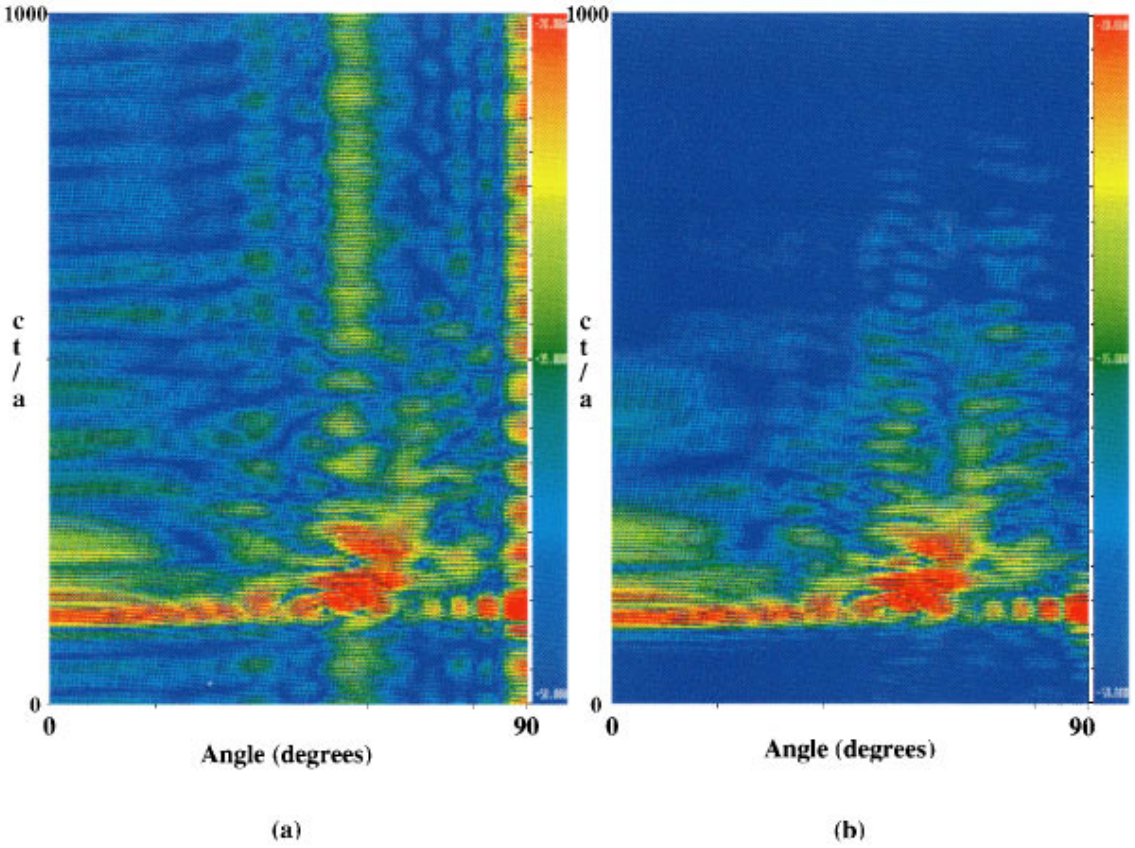


FIG. 7. Time domain response of cylinder for $2.5 \leqslant ka \leqslant 2.9$ in (a) shallow water with horizontal array of sources and vertical array of receivers and (b) free space.

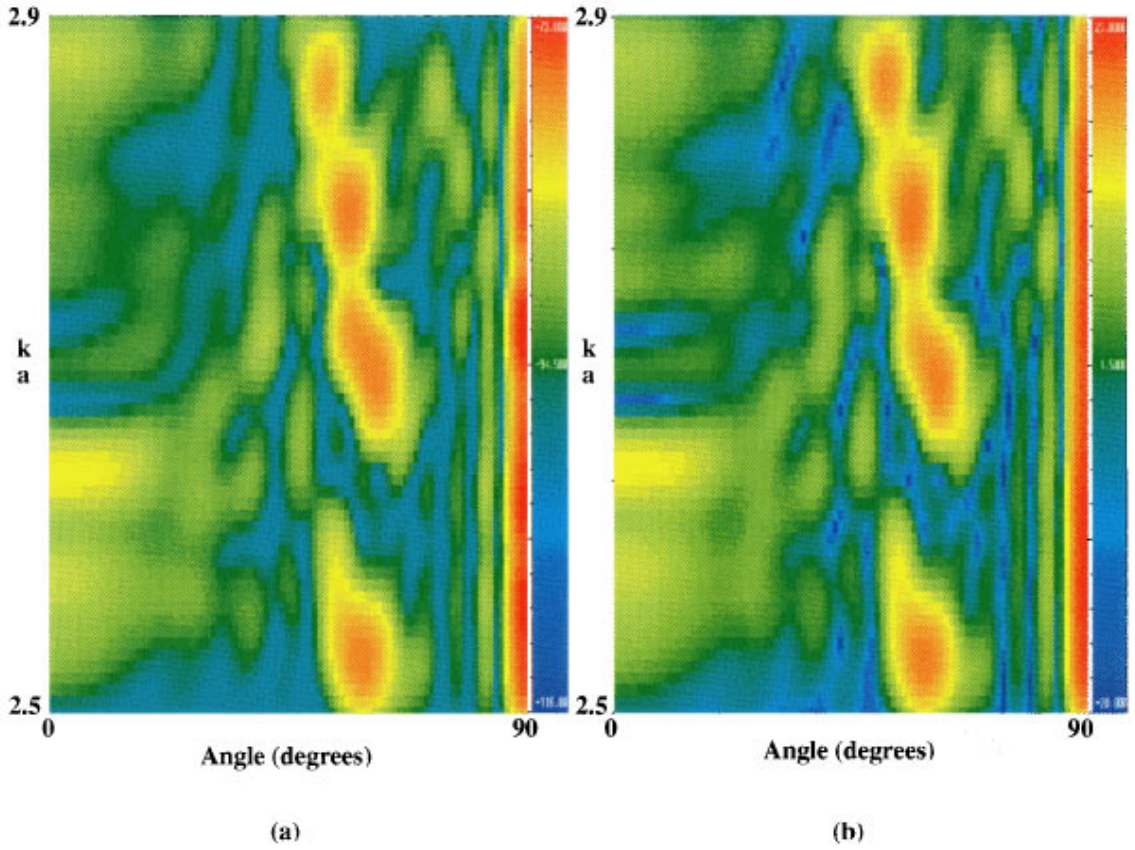


FIG. 8. Frequency domain response of cylinder in (a) shallow water with horizontal array of sources and vertical array of receivers and (b) free space.

ter environment, beamforming techniques that suppress low-order modes may be used in addition to temporal windowing techniques that remove from the received signal higher-order modes. Such a procedure is more feasible with vertical arrays because of the larger spread of wave numbers in the vertical direction. The smaller spread of wave numbers in the horizontal direction makes beamforming more difficult with horizontal arrays because of the condition of the matrices that require inversion in that case.

[1] T. C. Yang and T. W. Yates, ''Scattering from an object in a stratified medium. II. Extraction of scattering signature,'' J. Acoust. Soc. Am. **96**, 1020–1031 (1994).

[2] G. V. Frisk, *Ocean and Seabed Acoustics* (Prentice-Hall, Englewood Cliffs, NJ, 1994), pp. 111–116.

[3] A. Sarkissian, ''Multiple scattering effects when scattering from a target in shallow water,'' J. Acoust. Soc. Am. **96**, 3137–3144 (1994).

[4] E. P. McDaid, D. Gillette, and D. Barach, ''Interaction of a target with its acoustic environment,'' NRAD Technical Report 1679, September 1994.

[5] F. Ingenito, ''Scattering from an object in a stratified medium,'' J. Acoust. Soc. Am. **82**, 2051–2059 (1987).

[6] Yu. A. Kravtsov, V. M. Kuz'kin, and V. G. Petnikov, ''Wave diffraction by regular scatterers in multimode waveguides,'' Sov. Phys. Acoust. **30**, 199–202 (1984).

[7] B. B. Baker and E. T. Copson, *The Mathematical Theory of Huygens' Principle* (Oxford, New York, 1950), 2nd ed.

[8] A. Bostrom, ''Transmission and reflection of acoustic waves by an obstacle in a waveguide,'' Wave Motion **12**, 167–184 (1984).

[9] R. H. Hackman and G. S. Sammelmann, ''Multiple-scattering analysis for a target in an oceanic waveguide,'' J. Acoust. Soc. Am. **84**, 1813–1825 (1988).

[10] G. V. Norton and M. F. Werby, ''A numerical technique to describe acoustic scattering and propagation from an object in a waveguide,'' J. Appl. Phys. **70**, 4101–4112 (1991).

[11] T. W. Dawson and J. A. Fawcett, ''A boundary integral equation method for acoustic scattering in a waveguide with nonplanar surfaces,'' J. Acoust. Soc. Am. **87**, 1110–1125 (1990).

[12] A. Sarkissian, ''Method of superposition applied to scattering from a target in a bounded medium,'' J. Acoust. Soc. Am. **95**, 2340–2345 (1994).

[13] C. T. Tindle, K. M. Guthrie, G. E. J. Bold, M. D. Johns, D. Jones, K. O. Dixon, and T. G. Birdsall, ''Measurements of the frequency dependence of normal modes,'' J. Acoust. Soc. Am. **64**, 1178–1185 (1978).

[14] H. Alik, R. Dees, S. Moore, and D. Pan, ''Sara-2d User's Manual,'' version 95-3, BBN Systems and Technologies, New London, Ct, 1995.

[15] O. C. Zienkiewicz, C. Emson, and P. Bettess, ''A novel boundary infinite element,'' Int. J. Numer. Methods Eng. **19**, 393–404 (1983).

[16] O. C. Zienkiewicz, K. Bando, P. Bettess, C. Emson, and T. C. Chiam, ''Mapped infinite elements for exterior wave problems,'' Int. J. Numer. Methods Eng. **21**, 1229–1251 (1985).

[17] M. Porter, ''The KRAKEN Normal Mode Program,'' Naval Research Laboratory Report, NRL-92-6920 (1992).

Inversion of bottom/subbottom statistical parameters from acoustic backscatter data[a)]

Altan Turgut
Naval Research Laboratory, Acoustics Division, Washington, DC 20375

(Received 29 December 1996; accepted for publication 26 January 1997)

Inversion of statistical parameters of a bottom/subbottom scattering model is investigated by using genetic algorithms for both synthetic and real data. The bottom/subbottom scattering model used in the calculations is a modified version of Lyons, Anderson, and Dwan [J. Acoust. Soc. Am. **79**, 1410–1422 (1986)] in which correlation between subbottom density, compressibility, and sound-speed fluctuations is established through Wood's equation [*A Textbook of Sound* (Macmillan, New York, 1941)], and volume inhomogeneities are described by von Karman type autocorrelation functions [T. von Karman, J. Mar. Res. **7**, 252–264 (1948)]. The inversion is posed as an optimization problem which is solved (by a controlled Monte Carlo search using genetic algorithms) to find the optimum set of statistical parameters that minimizes the quadratic deviation between measured and calculated backscatter data. *A posteriori* probabilities calculated at the end of the search are used for error estimation and indication of relative importance of model parameters. This information helps to further assess the relative importance of two major scattering mechanisms due to bottom roughness and subbottom inhomogeneities. Such assessment is successfully demonstrated by using synthetic and real backscatter data for sandy, silty, and muddy sediments. Finally, inverted statistical parameters for a sandy site at Biscayne Bay, Florida confirm the results of simultaneous tomographic measurements indicating that scattering by subbottom inhomogeneities plays a minor role for this particular site. [S0001-4966(97)00908-9]

PACS numbers: 43.30.Ft, 43.30.Gv, 43.30.Ky, 43.30.Ma [JHM]

INTRODUCTION

During the past three decades, sound scattering from the ocean bottom has been investigated by several researchers, most of whom have emphasized the scattering from the water–sediment interface (e.g., Marsh, 1961; Kuo, 1963). However, in the case of a smooth bottom with soft sediments, their theoretical models predict much lower backscatter levels than those obtained by measurements. Excess levels of measured backscatter are usually attributed to volume scattering from subbottom inhomogeneities. Beginning in the early 1960's, several theoretical models for volume scattering have been developed. Stockhausen (1963) developed a volume scattering model which includes refraction and attenuation in a statistically homogeneous sediment with a flat sea bottom and small spherical scatterers. Crowther (1983) extended Kuo's model for interface scattering by introducing a complex sound speed in the sediment. He also developed a volume scattering model for a flat, refracting interface and compared his model results with previous experimental results for moderate (1.6–6.4 kHz) and very high (400–1000 kHz) frequencies. Jackson *et al.* (1986) developed a composite roughness model for bottom interface scattering and combined it with Stockhausen's model for volume scattering. Mourad and Jackson (1989) introduced an improved version of the composite roughness model and related the bottom scattering strength to a number of measurable bottom parameters. Tang (1991) has shown that measured spatial correlation of the scattered field might also be used to remotely determine the parameters critical to volume scattering. Jackson and Briggs (1992) were able show that the interface scattering model at a sandy site, and the volume scattering model at two silty sites, fit their backscattering data. In their model they have left the sediment volume scattering strength as a free parameter due to lack of statistical information on sediment volume inhomogeneities. Recently, Lyons *et al.* (1994) incorporated Stockhausen's volume scattering model including composite roughness with the addition of scattering from subbottom interfaces. They showed that high/low backscatter levels of GLORIA imagery are due to the presence/absence of an inhomogeneous silt–clay layer overlying an acoustically hard sand layer. Similar to Crowther's (1983) treatment, they described the sediment volume inhomogeneities using an anisotropic exponential correlation function and obtained some of the model parameters from ground truth core data (one-dimensional) leaving the horizontal correlation length as a free parameter. Turgut (1991) introduced a tomographic method for measuring sound-speed inhomogeneities in a vertical cross section of sediment so that a two-dimensional correlation function can be estimated. The two-dimensional correlation function provides information on both vertical and horizontal correlation lengths, as well as layer dipping at certain azimuthal directions. Also, several authors have attempted to estimate parameters of interface scattering models from backscatter data. De Moustier and Alexandrou (1990) implemented a qualitative curve-fitting method to estimate two parameters of the Helmholtz–Kirchoff (H–K) model from their multi-beam echo-sounder data. Michalopoulou and Alexandrou (1996) used a Bayesian

[a)]Experimental part of this work was completed while the author was at the University of Miami.

approach with a maximum likelihood estimation method to estimate the H–K model parameters from the same data set. Bourgeois and Caruthers (1992) implemented backpropagation neural networks to estimate six parameters of their bistatic scattering strength model with partial success. Similarly, Gott and Martinez (1993) reported limited success in estimating four parameters of both Rayleigh–Rice (R–R) and H–K models using neural networks.

In this paper, both interface roughness parameters of a R–R model and subbottom volume scattering parameters of a modified Stockhausen model are simultaneously inverted by using parallel genetic algorithms (GAs) in a global optimization method. Along with successful inversions of both synthetic and real data, *a posteriori* probabilities are estimated and the relative importance of each parameter (interface or volume) is determined for a given data set.

The paper is organized as follows: In Sec. I, after a brief definition of interface and volume scattering cross sections, 3D sound-speed structure in the sediment is described by a von Karman-type spectrum. Correlation among sediment density, compressibility, and sound-speed fluctuations is also established using Wood's equation so that volume backscattering strength can be defined by the variance of a single parameter (sound speed). In Sec. II, a standard GA is briefly described and its extension to parallel GAs used in the inversion is outlined. In Sec. III, sensitivity analysis and numerical examples of inversion from synthetic data for muddy, silty, and sandy sediments are given. In Sec. IV, tomographic measurements of sediment inhomogeneities at the Biscayne Bay site are briefly outlined. Then, volume scattering parameters estimated by the tomographic measurements are compared with those of GA inversion from backscatter data. Finally, backscatter data from Arafura Sea and Quinault Range sites (Jackson and Briggs, 1992) are inverted. Conclusions are given in Sec. V. For completeness, the composite roughness model of Jackson *et al.* (1986) and volume scattering model of Lyons *et al.* (1994) with modifications are briefly outlined in the Appendix.

I. BOTTOM BACKSCATTERING MODEL

Bottom backscattering strength, $S_b(\theta)$, is defined as the decibel equivalent of the sum of interface and volume scattering cross sections (Urick, 1983) as

$$S_b(\theta) = 10 \log_{10}[\sigma_{\rm cr}(\theta) + \sigma_v(\theta)], \tag{1}$$

where θ is the grazing angle, and $\sigma_{\rm cr}(\theta)$ and $\sigma_v(\theta)$ are dimensionless backscattering cross sections per unit solid angle per unit area due to interface roughness and volume scattering, respectively. Both cross sections are calculated using perturbation approximations, and corrections for shadowing and large-scale bottom slope are included in the context of the composite roughness approximation (McDaniel and Gorman, 1983) as

$$\sigma_{\rm cr}(\theta) = S(\theta,s) F(\theta,\sigma_{\rm pr},s), \tag{2}$$

and

$$\sigma_v(\theta) = S(\theta,s) F(\theta,\sigma_{\rm pv},s), \tag{3}$$

where $S(\theta,s)$ is the shadowing correction, $F(\theta,\sigma_{\rm pr},s)$ is a slope averaging integral, s is the rms slope, $\sigma_{\rm pr}$ is the interface backscattering cross section in the Rayleigh–Rice perturbation approximation, and $\sigma_{\rm pv}$ is the volume backscattering cross section in the Born perturbation approximation. A detailed description of each cross section is given in the Appendix [Eqs. (A4) and (A18)].

A. Small-scale sediment sound-speed structure as a 3-D random field

Sound-speed structure in the sediment is rather complex so that a determistic description of such a field is almost impossible for a given site considering the length scales of interest (a few centimeters to tens of meters). Stochastic description and parameterization of the field in terms of its statistical properties is more useful for an acoustic backscattering model. We describe the sediment sound-speed structure by partitioning the field into a deterministic part $c_0(\mathbf{x})$ and a stochastic part $\Delta c(\mathbf{x})$ as

$$c(\mathbf{x}) = c_0(\mathbf{x}) + \Delta c(\mathbf{x}), \tag{4}$$

where $\mathbf{x}$ is the coordinate vector. The deterministic part $c_0(\mathbf{x})$ represents site-specific large-scale features of a given geological province. The stochastic part $\Delta c(\mathbf{x})$ represents small-scale sound-speed structure which is modeled as a zero-mean quasi-stationary random process. Second-order statistics of the field described by a 3D power spectrum are needed to model the acoustic scattering from such a medium. Assuming $c_0(\mathbf{x})$ is constant within the sediment volume, derivation of a volume scattering model is given in the Appendix [see Eqs. (A5)–(A16)]. Several forms of spectral representations of volume inhomogeneities were proposed by previous researchers, with certain spectral parameters to be estimated from available sediment core samples. Based on their limited number of core measurements, Crowther (1983) and Lyons *et al.* (1994) suggested a 3-D exponential correlation function for the volume inhomogeneities corresponding to a 3-D power spectrum in the form:

$$S_e(\xi) = \frac{a^2 b \mu}{\pi^2 (1+\xi^2)^2}, \tag{5}$$

where $\xi = [(a^2(\xi_x^2+\xi_y^2) + b^2\xi_z^2]^{1/2}$ is the magnitude of the wave number vector $\boldsymbol{\xi}$, ξ_x and ξ_y are the horizontal components of $\boldsymbol{\xi}$, ξ_z is the vertical component of $\boldsymbol{\xi}$, a is the horizontal scale factor, b is the vertical scale factor, and μ is the sound-speed variance.

At large wave numbers ($a\xi_x, a\xi_y \gg 1$, and $b\xi_z \gg 1$) the above spectrum rolls off at the constant rate ξ^{-4}. However, Briggs' (1994) estimation of correlation functions from the cores taken at 14 different sites showed distinct departures from an exponential correlation function, suggesting the need for an arbitrary roll-off rate. A pure power-law spectrum, proposed by Tappert (1991), includes such an arbitrary roll-off rate with the form

$$S_p(\xi) = \frac{\beta_1 a^2 b^{\beta_1} B}{2\pi}\, \xi^{-(\beta_1+2)}, \tag{6}$$

where B is the structure constant, and β_1 is the one-dimensional spectral exponent. The parameters B and β_1 are usually obtained from vertical core sampling assuming horizontal layering. In the case of layer dipping, they should be modified to represent the statistics along the axis normal to dipped layering. Although S_p generalizes S_e to an arbitrary asymptotic roll-off rate, it is inadequate at low wave numbers, since extrapolating a pure power law too far in this direction leads to unrealistically large values of sound-speed variances. A more adequate spectrum of the von Karman-type is proposed to describe sediment sound-speed inhomogeneities including horizontal anisotropy as

$$S_v(\xi)=\mu a_x a_y b\,\pi^{-3/2}\,\frac{\Gamma(m)}{\Gamma(m-3/2)}\,(1+\xi^2)^{-m},\quad m>3/2, \tag{7}$$

where a_x and a_y are the horizontal scale factors ($a_x \neq a_y$), Γ is the Gamma function, and m is the volume spectral exponent. The above spectrum is ellipsoidal since the components of the wave number vector $\boldsymbol{\xi}$ are scaled by direction-dependent outer scales ($|\boldsymbol{\xi}|=\xi=[(a_x\xi_x)^2+(a_y\xi_y)^2+(b\xi_z)^2]^{1/2}$). Similarly, the associated correlation function $B_v(r)$ is also ellipsoidal since it depends on the distance $r(\mathbf{x})=[(x/a_x)^2+(y/a_y)^2+(z/b)^2]^{1/2}$ as

$$B_v(r)=\frac{\mu}{2^{p-1}\Gamma(p)}\,r^p K_p(r), \tag{8}$$

where $K_p(r)$ is Bessel's function of the second kind of order $p=m-3/2$ (Tatarskii, 1961).

The 2-D wave number spectrum $S(\xi_x,\xi_z)$ is obtained by integrating the 3-D spectrum over the wave number ξ_y:

$$S_v(\xi_x,\xi_z)=\mu a_x b\,\pi^{-1}\,\frac{\Gamma(m-1/2)}{\Gamma(m-3/2)}\times(1+a_x^2\xi_x^2+b^2\xi_z^2)^{-m+1/2}. \tag{9}$$

Estimation of the above spectral parameters μ, a_x, b, and m was attempted first by Turgut (1991) using high-resolution tomographic measurements of sound-speed inhomogeneities. Further integration of Eq. (9) over the wave number ξ_x yields a 1-D spectrum:

$$S_v(\xi_z)=\mu b\,\pi^{-1/2}\,\frac{\Gamma(m-1)}{\Gamma(m-3/2)}\,(1+b^2\xi_z^2)^{-m+1}, \tag{10}$$

in which the spectral parameters μ, b, and m can be obtained from sediment core analysis.

B. Correlation of sediment compressibility, density, and sound-speed fluctuations

In this section, relations among sediment density, compressibility, and sound-speed fluctuations are derived, and their effects on the backscattering strength are analyzed. In the Appendix, scattering amplitude $f(\hat{\mathbf{0}},\hat{\mathbf{i}})$ in Eq. (A8) is defined in terms of relative compressibility and density fluctuations ($\delta\kappa/\kappa_0$ and $\delta\rho/\rho_0$) around their mean values κ_0 and ρ_0, respectively. Correlation of compressibility and density fluctuations to sound-speed fluctuations enables us to represent the field by the fluctuations of a single parameter (sound speed) in the sediment. Sediment compressibility and density fluctuations can be calculated from the sound-speed fluctuations, assuming that Wood's (1941) equation,

$$\kappa_0=(1-n)\kappa_r+n\kappa_f, \tag{11}$$

is valid for most marine sediments, where n is the sediment porosity, κ_r is the compressibility of the grain, and κ_f is the compressibility of the pore fluid. Similarly, bulk density of the sediment is defined as

$$\rho_0=(1-n)\rho_r+n\rho_f, \tag{12}$$

where ρ_r is the density of the grain and ρ_f is the density of the pore fluid. Relative sound-speed fluctuation in the medium is related to density and compressibility fluctuations, in first order, as

$$\frac{\delta c}{c_0}=-\frac{1}{2}\left(\frac{\delta\kappa}{\kappa_0}+\frac{\delta\rho}{\rho_0}\right), \tag{13}$$

where $c_0=1/(\kappa_0\rho_0)^{1/2}$ is the spatial average of sound speed. Then, density and compressibility fluctuations can be expressed in terms of sound-speed fluctuations as

$$\frac{\delta\rho}{\rho_0}=-2\left(\frac{\gamma_\rho-1}{\gamma_\rho+\gamma_\kappa-2}\right)\frac{\delta c}{c_0}, \tag{14}$$

and

$$\frac{\delta\kappa}{\kappa_0}=-2\left(\frac{\gamma_\kappa-1}{\gamma_\rho+\gamma_\kappa-2}\right)\frac{\delta c}{c_0}, \tag{15}$$

where $\gamma_\rho=\rho_r/\rho_0$ and $\gamma_\kappa=\kappa_r/\kappa_0$. For a measured mean value of sound speed $c_0=1650$ m/s at the Biscayne Bay site, calculated porosity, mean density, mean compressibility, and ratios of their variation to sound-speed variation are given in Table I. For backscattering calculations, considering the opposite directions of incoming and scattered waves [$\cos(\hat{\mathbf{0}},\hat{\mathbf{i}})=-1$], the kernel of the integral in Eq. (A8) can be related to the sound-speed fluctuations as

$$\frac{\delta\kappa}{\kappa_0}-\frac{\delta\rho}{\rho_0}=V_{\kappa\rho}^{(B)}\,\frac{\delta c}{c_0}, \tag{16}$$

where the parameter $V_{\kappa\rho}^{(B)}$ is given as $2(\gamma_\rho-\gamma_\kappa)/(\gamma_\rho+\gamma_\kappa-2)$. In Fig. 1, the ratio of relative density fluctuations to sound-speed fluctuations $(\delta\rho/\rho_0)/(\delta c/c_0)$, the ratio of relative compressibility fluctuations to the sound-speed fluctuations $(\delta\kappa/\kappa_0)/(\delta c/c_0)$, and the parameter $V_{\kappa\rho}^{(B)}$ are plotted as a function of mean sound speed c_0. Notice that the square of parameter $V_{\kappa\rho}^{(B)}$ becomes significantly large for softer sediments ($c_0<1650$ m/s). Since the volume scattering cross section, Υ, is linearly dependent on the square of parameter $V_{\kappa\rho}^{(B)}$ [see Eq. (A15)], softer sediments should exhibit larger volume scattering strengths among the different types of sediments having the same sound-speed variance.

II. GENETIC ALGORITHMS

Genetic algorithms (GAs) are search algorithms based on natural selection and natural genetics (Goldberg, 1989). Since a detailed description of GAs and their application to geophysical (Stoffa and Sen, 1991) and geoacoustic (Gerstoft, 1993) parameter estimation are given elsewhere, a brief outline is presented in this section. In a standard GA, ran-

TABLE I. Physical properties used for Biscayne Bay sediments.

Physical properties	Symbols (units)	Assumed values
Compressibility of grain	κ_r $(\mathrm{N/m^2})^{-1}$	2.5×10^{-11}
Compressibility of the pore fluid	κ_f $(\mathrm{N/m^2})^{-1}$	4.35×10^{-10}
Density of the grain	ρ_r $(\mathrm{kg/m^3})$	2700
Density of the pore fluid	ρ_f $(\mathrm{kg/m^3})$	1024

Calculated properties for Biscayne Bay sediments ($c_0=1650$ m/s)

Physical properties	Symbols (units)	Calculated values
Porosity	n	0.38
Mean compressibility of sediment	κ_0 $(\mathrm{N/m^2})^{-1}$	2.5×10^{-11}
Mean density of sediment	ρ_0 $(\mathrm{kg/m^3})$	2024
Variability ratio of density to sound speed	$(\delta\rho/\rho_0)/(\delta c/c_0)$	1.3
Variability ratio of compressibility to sound speed	$(\delta\kappa/\kappa_0)/(\delta c/c_0)$	-3.3

domly selected values within a specified range are assigned to each model parameter. Then, model parameters are binary coded and formed into a string of haploid chromosomes. A specified number of chromosomes is generated to obtain a finite population on which three operations, namely: reproduction, crossover, and mutation are applied at each generation step. Reproduction is a process that selects the models according to their objective functions from a randomly chosen population of models **m**. This means that models producing optimal values of objective function $\phi(\mathbf{m})$ will have a higher probability of contributing one or more offspring in the next generation. After reproduction, a mating pool is established and crossover is performed between randomly paired models in the mating pool. Crossover is simply partial exchange of bits between the paired models. This structured yet randomized information exchange creates new models with a minimum disruption to selection strategy dictated by reproduction alone. It is subject to a specified crossover probability, p_x. High p_x means most of the pairs will mate and produce offspring. In case of low p_x, most of the parents will be carried on to the next generation without mating. In our application, a multi-point crossover was used by simultaneously exchanging information between corresponding model parameters within the paired model vectors. Mutation is necessary to ensure diversity within the population and to prevent the search by reproduction and crossover from being trapped around a local minimum. A randomly selected bit in the parameter code is altered according to a specified mutation probability p_m. An update probability, p_u, is also used to control the influence of fit models from previous generations in the current generation. After the crossover, fitness functions of new models are calculated. For each new model, a model from the previous generation which was not selected for the mating pool is assigned at random. Objective functions of both the new model and the randomly selected old model are compared. If the objective function of the new model is better, the new model is always kept. If the objective function of the randomly selected old model is better, the new model is replaced by the old model with the specified update probability, p_u. We should point out that in our application, the exact shape of the objective function is not known. It might be multimodal due to the physics of the forward problem or overparameterization of the inverse problem. Goldberg and Richardson (1987) have shown that standard GAs with the above three operators (p_x, p_m, and p_u) converge to a single minimum among the several minima with nearly equal heights (called *genetic drift* by the geneticists). Several techniques such as *sharing* (Goldberg and Richardson, 1987) and *crowding* (De Jong, 1975), and their improved versions (Mahfoud, 1992; Miller and Shaw, 1995) have reported satisfactory results dealing with multimodal objective functions. As in Sen and Stoffa (1992), we used parallel GAs to circumvent the problem of *genetic drift*.

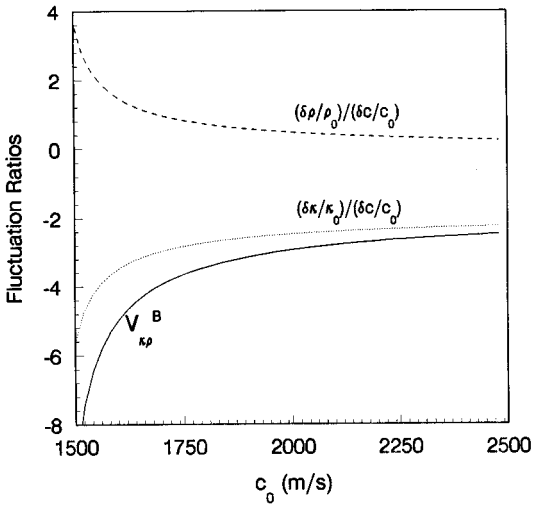


FIG. 1. Density and compressibility variations with respect to sound-speed variation, and parameter $V_{\kappa\rho}^{(B)}$ as a function of mean sound-speed c_0.

A minimization problem is formulated by calculating an objective, or cost function as the quadratic deviation between measured and calculated backscattering strengths,

$$\phi(\mathbf{m}^i)=\frac{\|\mathbf{S}_{\mathrm{obs}}-\mathbf{S}_{\mathrm{cal}}(\mathbf{m}^i)\|^2}{\|\mathbf{S}_{\mathrm{obs}}\|^2}, \tag{17}$$

where $\mathbf{S}_{\mathrm{obs}}$ and $\mathbf{S}_{\mathrm{cal}}(\mathbf{m}^i)$ are vectors of observed and calculated backscattering strengths (both interface and volume scattering components) as a function of grazing angle, and subscript i refers to the ith model vector within a randomly selected initial population. Here, $\|\cdot\|$ is the l^2 norm of a vector $[\|\mathbf{x}\|=(x_1^2+\cdots+x_n^2)^{1/2}]$.

An exponential probability of selection is defined relating the likelihood of selection to their cost functions as

$$P_s(\mathbf{m}^i) = \frac{e^{-\phi(\mathbf{m}^i)/T}}{\sum_{j=1}^{N} e^{-\phi(\mathbf{m}^j)/T}}, \tag{18}$$

where $P_s(\mathbf{m}^i)$ is the probability of selection for the ith model, N is the population size, and T is a control parameter. The control parameter T provides a controlled rate of convergence to a global minimum. It can be analogous to a cooling temperature used in the simulated annealing method described by Sen and Stoffa (1992). Alternatively, Sambridge and Drijkoningen (1992) used the standard deviation of the distribution, and Gerstoft (1993) used the minimum value of the distribution, both with satisfactory rates of convergence. In the current work, an effective cooling schedule was adapted from Sen and Stoffa (1992) in which the temperature was changed in a step-function fashion as follows:

$$\begin{aligned} T_k &= (0.89)^k \quad \text{for generations} \leqslant 5, \\ T_k &= 0.01(0.89)^k \quad \text{for generations} > 5, \end{aligned} \tag{19}$$

where k is the generation number. At the end of several parallel runs, the *a posteriori* probability density (PPD), $\sigma(\mathbf{m})$, the mean model parameters, $\langle \mathbf{m} \rangle$, and the *a posteriori* model covariance matrix, $\mathbf{C}_M$, are calculated as (Tarantola, 1987)

$$\sigma(\mathbf{m}) = \frac{e^{-\phi(\mathbf{m})/T}}{\sum e^{-\phi(\mathbf{m})/T}}, \tag{20}$$

$$\langle \mathbf{m} \rangle = \sum \mathbf{m}\sigma(\mathbf{m}), \tag{21}$$

$$\mathbf{C}_M = \sum \mathbf{m}\mathbf{m}^{\dagger}\sigma(\mathbf{m}) - \langle \mathbf{m} \rangle \langle \mathbf{m} \rangle^{\dagger}, \tag{22}$$

where T is the temperature and $\mathbf{m}$ is the model derived at that temperature, and the sum is taken over all the derived models during the parallel search. During each individual run, the above quantities were accumulated to estimate the PPD, mean, and covariance of the model parameters. Notice that the PPD is defined as the Gibbs probability distribution with a temperature T so that higher probabilities are assigned to the models picked during the later generations. The PPDs are good indicators of importance and uniqueness of each parameter. The square roots of the diagonal terms of the *a posteriori* model covariance matrix are the standard deviations of the model parameters from the mean, which is a measure of the accuracy of the inversion. The off-diagonal terms are the correlation coefficients which show the interaction between different model parameters.

III. INVERSION OF SURFACE AND VOLUME SCATTERING PARAMETERS

A. Sensitivity analysis

Modeling of both surface and volume backscattering requires as many as nine parameters, which can be grouped as surface, volume, and geoacoustic parameters. With the assumption of uncorrelated surface and volume parameters, the spectral exponent γ and spectral strength β in the surface roughness spectrum in Eq. (A1) determine surface scattering. On the other hand, volume scattering depends on the vertical scale factor b, the aspect ratio α (ratio of horizontal scale factor to vertical scale factor assuming $a_x = a_y$), the volume spectral exponent m, and the sediment sound-speed variance μ. The loss parameter δ, sound-speed ratio ν, and density ratio q (geoacoustic parameters) are expected to contribute to both surface and volume scattering. A sensitivity analysis is described below for the cases of muddy, silty, and sandy sediments to demonstrate the effect of each parameter on the backscattering strength. Backscattering strengths were calculated for the lower and upper bounds of each parameter, while keeping the other parameters constant (see Table II). Numerical values of surface parameters were selected based on the measurements of Jackson and Briggs (1992), and those of volumetric parameters were selected based on our tomographic measurements and the core analyses of Briggs (1994). Bounds of each parameter are needed to be known as *a priori* information so that a more effective search can be performed within physically plausible limits. Figure 2(a)–(c) shows dependence of backscattering strengths on each individual parameter for muddy, silty, and sandy sediments for the lower (solid line) and upper (dotted line) bounds. As expected, Fig. 2(a) shows relatively stronger dependency of backscattering strength on volume and geoacoustic parameters (μ, α, m, b, δ, ν, and q) than that on surface parameters (γ and β) in muddy sediments. Figure 2(c) depicts a reversed dependency of backscattering strength for sandy sediments. In this case, effects of surface parameters (γ and β) are much more significant than those of the volume parameters (α, b, and m). Figure 2(b) shows some degree of dependence of backscattering strength on all the parameters for silty sediments. Considering a typical experimental error in backscattering measurements to be as much as 3 dB (Jackson and Briggs, 1992), inversion of certain parameters having minor effects on the backscattering strength will be almost impossible from the real data. In our

TABLE II. True model parameters and their bounds used in the GA inversion.

		Sediment type			Bounds	
Model parameters	Symbols	Muddy	Silty	Sandy	Lower	Upper
Surface spectral exponent	γ	3.2	3.5	3.8	2.9	3.9
Surface spectral strength	β	0.0005	0.003	0.005	0.000 01	0.01
Loss parameter	δ	0.005	0.01	0.015	0.002	0.02
Vertical scale factor	b	0.1	0.15	0.2	0.01	0.5
Aspect ratio	α	10.0	7.5	6.0	1.0	15.0
Sound-speed variance	μ	0.005	0.001	0.0005	0.0001	0.01
Sound-speed ratio	ν	0.99	1.01	1.1	0.98	1.3
Density ratio	q	1.55	1.7	1.85	1.4	2.1
Volume spectral exponent	m	1.6	2.0	2.5	1.5	3.0

parallel GA inversion of both noise-free synthetic data and real data, *a posteriori* probabilities of these parameters are also calculated to investigate the physical importance of each individual parameter for the case of muddy, silty, or sandy sediments. Results of these calculations are given in Sec. III B.

Next, we investigate the objective function in Eq. (17) by changing three selected parameters within their specified bounds. All the other parameters are kept constant since a visual display of the objective function is not possible for higher dimensions. Figure 3(a) shows objective functions as a function of three volume parameters (b, μ, and m) for muddy, silty, and sandy sediments. Figure 3(b) shows objective functions as a function of one geoacoustic (ν) and two surface (γ and β) parameters for the same types of sediments. The objective function is minimum when the parameters take their true values, which are given in Table II. Each three-parameter objective function depicts a large and complicated shape of a minimum region arbitrarily defined as $\phi(\mathbf{m}) < 10^{-6}$. The minimum region shows mostly unimodal behavior for most of the parameters, except for the volume spectral exponent m. Multimodal behavior and some local minima might also exist for the other parameters as we increase the dimension of the objective function. In most cases of inverting noise-free synthetic data, a value of $5.\times 10^{-8}$ was achieved for the objective function at the end of 80 generations. This value was about $5.\times 10^{-6}$ for the real data due to other physical effects and measurement errors which are not included in our backscattering model. The minimum region of the objective function has a complicated shape for silty sediments. For sandy sediments, the objective functions calculated by changing volume parameters show a rather large volume of the minimum region, indicating the insensitivity to the volume parameters α, b, and m. On the other hand, strong dependency of the objective function on the surface parameters (β and γ) is clearly shown for the sandy sediment. Similar behavior of the objective function is observed for the sensitivity of the muddy sediment on the surface parameters in which isosurface planes being parallel to the γ-β plane indicate insensitivity to the surface parameters γ and β.

Finally, we explore the frequency dependency of our backscattering model. An exponential correlation function of sound-speed and density fluctuations obtained from core data taken from several geological provinces (Crowther, 1983; Lyons, 1994) and tomographic measurements at the Biscayne Bay site favors the use of spectrum $S_e(\xi)$ in Eq. (5) (a special case of von Karman spectrum with the spectral exponent $m=2$). However, Briggs' (1994) sediment core analyses from 14 different sites showed some variability in the correlation functions, indicating the need for a more general model. Unfortunately, both of these ''ground-truth'' measurements (core sampling and tomography) are band limited (limited by the sampling interval and area of coverage). A much wider spectrum of length scales might be needed to describe the angle-dependent backscattering

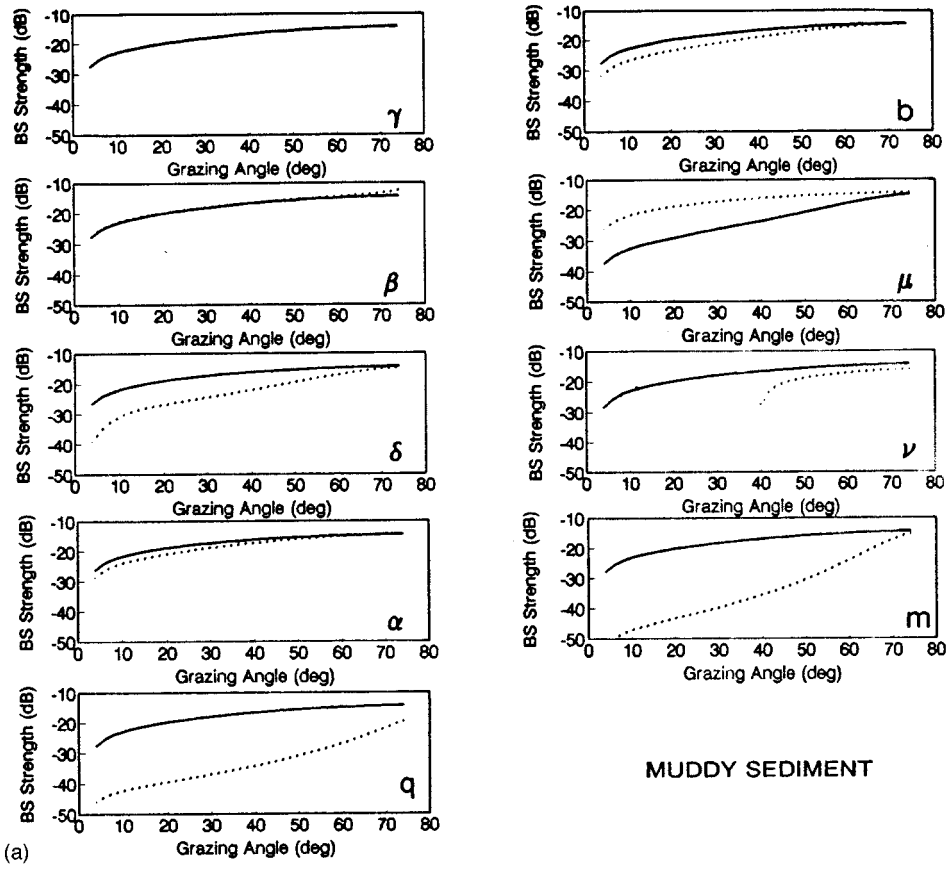
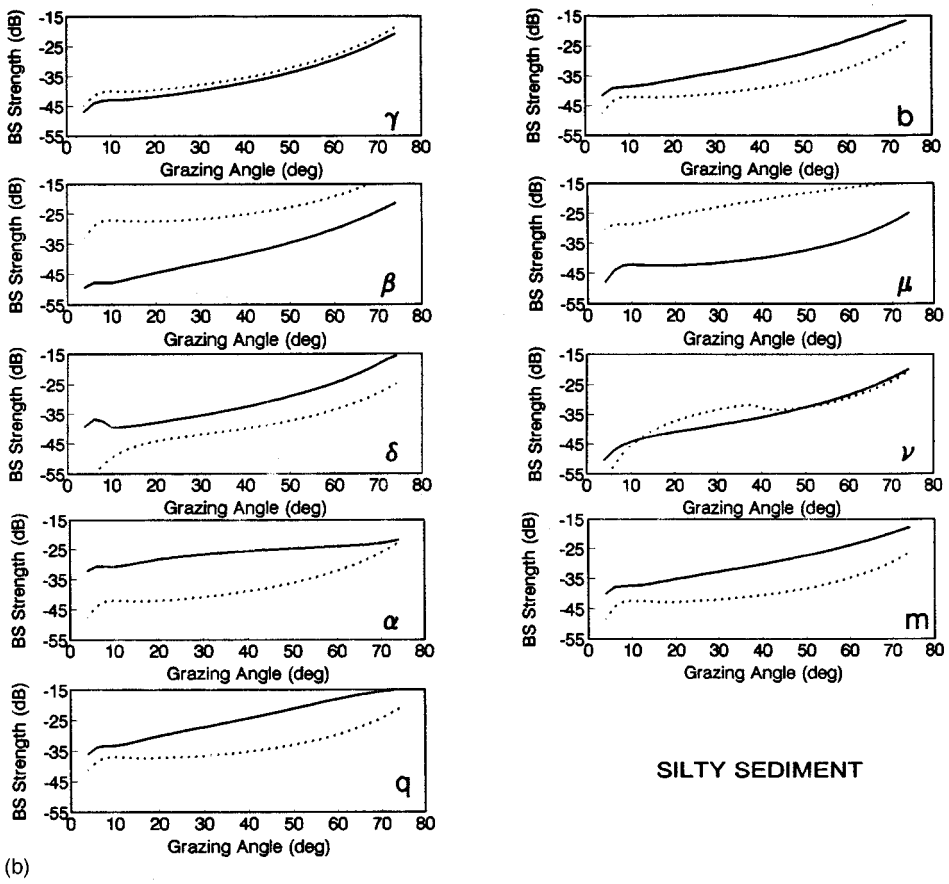
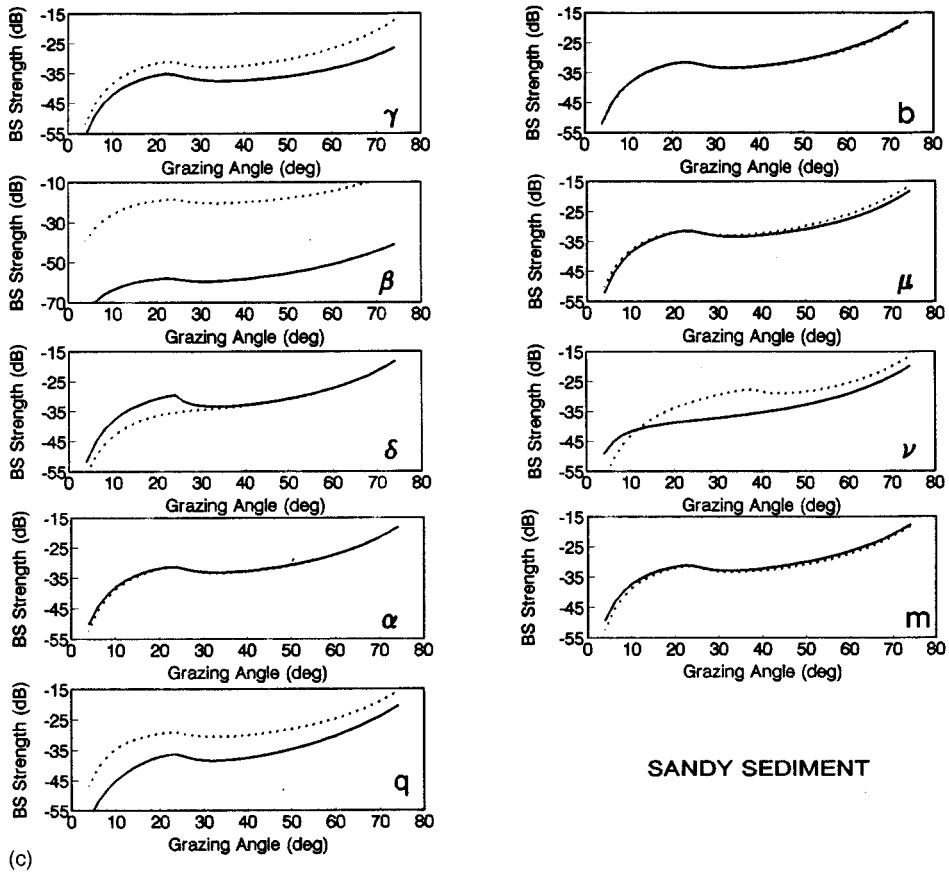


FIG. 2. Backscattering strength levels for bounds of each model parameter for (a) muddy, (b) silty, and (c) sandy sediments (solid line is the lower bound and dotted line is the upper bound).



(b)



(c)

FIG. 2. (*Continued.*)

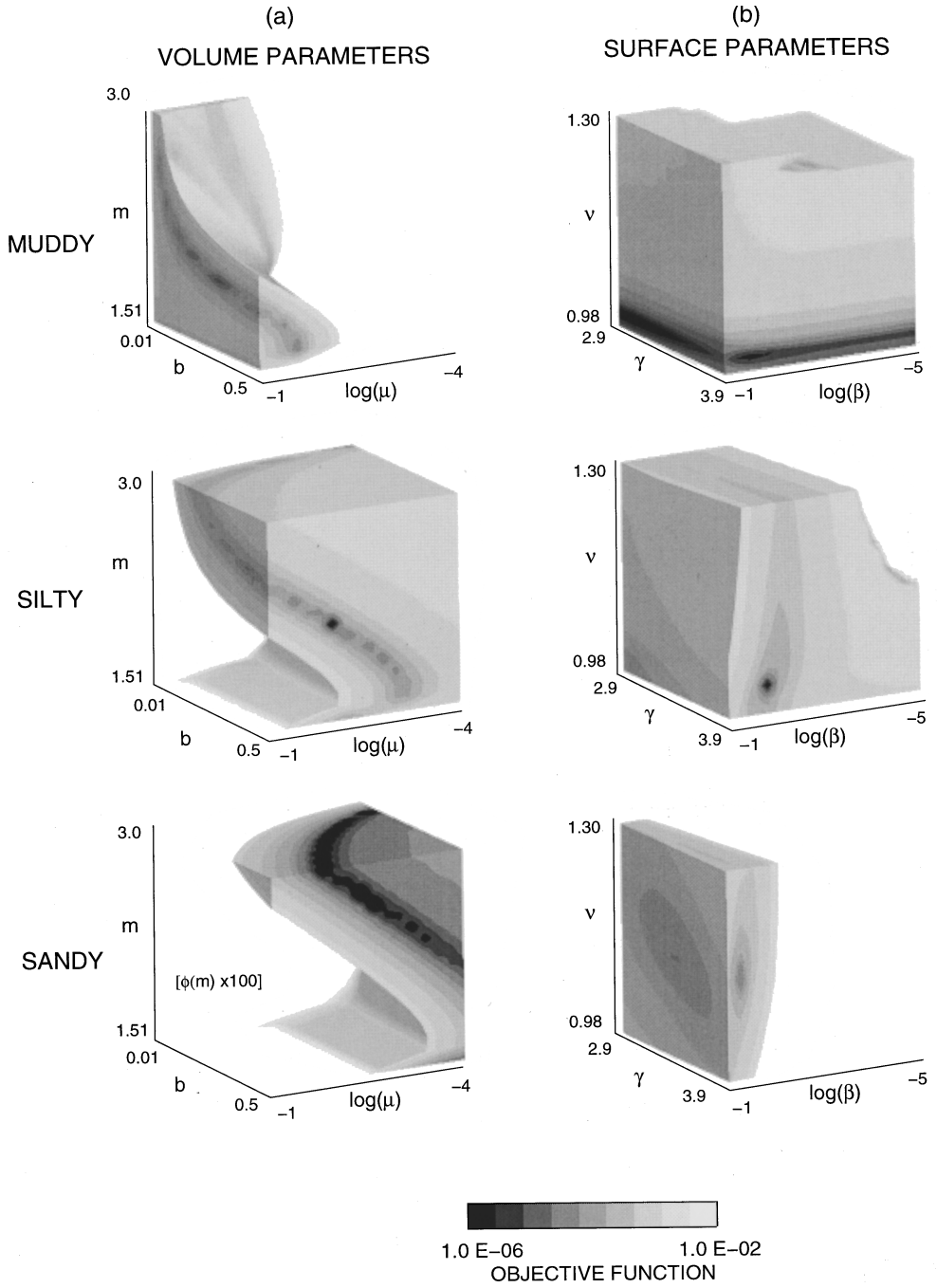


FIG. 3. Three-dimensional display of the objective functions as a function of (a) volume (b, m, and μ) and (b) geoacoustic (ν) and surface (γ and β) parameters for muddy, silty, and sandy sediments.

strength even for single-frequency measurements. Therefore, one should be careful to draw any conclusion on the frequency dependency of the backscattering from such band-limited ''ground-truth'' measurements. In Fig. 4, volume backscattering strength is plotted as a function of grazing angle and frequency for muddy sediments having spectral exponent values $m=1.51$, 2.0, and 2.5, respectively. As expected, the volume backscattering strength is almost constant for $m=1.51$ with the assumption of linear frequency-dependent intrinsic attenuation in the sediment [see Eqs. (A15) and (A16)]. The model predicts an increase in volume backscattering strength from 100 Hz to 1 kHz for muddy sediments with the spectral exponential values of $m=2.0$ and $m=2.5$. In this region, frequency dependency of volume backscattering is controlled by all the volume parameters (b, α, and m) in a complicated fashion. In the high-frequency region (between 1 kHz and 100 kHz), the volume backscattering strength decreases with the frequency for $m=2.0$ and $m=2.5$, with a higher rate for the latter case ($m=2.5$). The frequency dependency of the volume backscattering is mainly determined by the volume spectral exponent m in the high-frequency region.

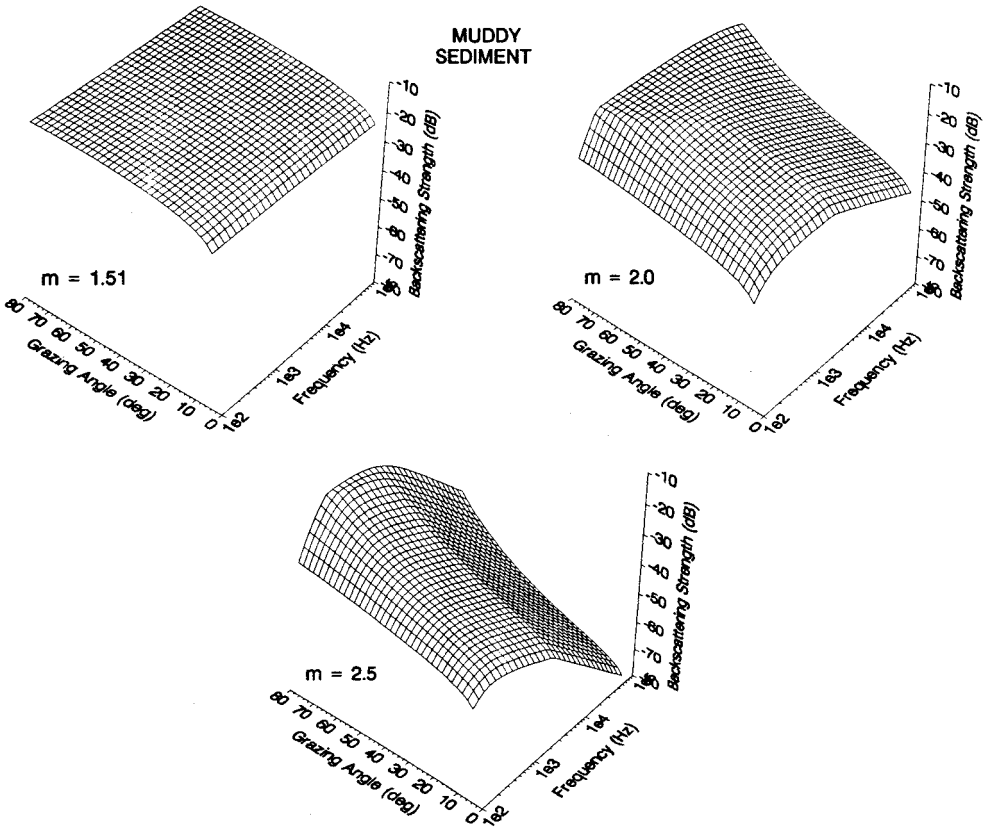


FIG. 4. Frequency and grazing angle dependency of volume backscattering strength for muddy sediments with different volume spectral exponent values ($m=1.51$, 2.0, and 2.5).

B. Inversion of synthetic backscattering data

The genetic algorithms described in Sec. II were implemented to invert the statistical and geoacoustic parameters (γ, β, b, α, μ, δ, ν, q, and m) using synthetic backscattering-strength data as a function of grazing angle. The synthetic data were generated at a frequency of 5 kHz for three sets of parameters corresponding to sandy, silty, and muddy sediments. In addition to specified limits of each parameter, additional constraints can be introduced by incorporating data from core samples, roughness measurements, and tomographic measurements, or by implementing additional theories to establish relations between certain parameters. Interface backscattering strengths are calculated using a composite roughness model (Jackson *et al.*, 1986) which is outlined in the Appendix. This model assumes a power-law dependency of the roughness spectrum, includes both small-scale roughness height and large-scale slope of the bottom, and is valid for moderate grazing angles (<70 deg). An outline of volume backscattering theory used in the model is also given in the Appendix. For all the cases, inversion was started with a population size (number of models) of $N=100$ and an initial temperature of $T_0=1$. Crossover, mutation, and update probabilities were also kept the same for all the cases ($p_x=0.6$, $p_m=0.001$, and $p_u=0.9$). An eight-bit chromosome length was used for each of the nine parameters defining a search space of $(2^8)^9\approx 5\times 10^{21}$ models. Convergence to the minimum objective function was achieved in 80 generations which required 8000 forward calculations. Synthetic backscattering data for grazing angles from 10 to 70 deg with 2-deg increments were used in all the inversions. Figure 5(a)–(c) shows PPDs of each model parameter (γ, β, δ, b, α, μ, ν, q, and m) for muddy, silty, and sandy sediments, respectively. True values of each parameter are also plotted as a vertical (dotted) line. In general, geoacoustic parameters (especially sound-speed ratio) are well estimated in all cases. Estimation is relatively better for surface parameters in sandy sediments, for volume parameters in muddy sediments, and for all of the parameters in silty sediments. The mean and one standard deviation of the parameters for sandy, silty, and muddy sediments were estimated after 20 parallel GA runs using different starting populations, and are shown in Fig. 5(a)–(c) also. In general, true values fall within one standard deviation which was calculated from the diagonal elements of *a posteriori* covariance matrix. This is reassuring especially considering the number of model parameters inverted simultaneously for different types of sediments in which parameter dependency of the cost function changes dramatically for each case.

During the search of model space, use of unrealistic density ratio and sound-speed ratio value pairs increases the complexity of inversion since these are the parameters in both surface and volume calculations [see Eqs. (A3), (A10), and (A17)]. Using Eq. (11) and Eq. (12), we establish a correlation between density and sound speed to reduce the number of parameters from nine to eight. By doing that, we eliminate the process of searching physically unrealistic portions of the model space (e.g., very high sound speeds for muddy sediment or very low densities for sandy sediments,

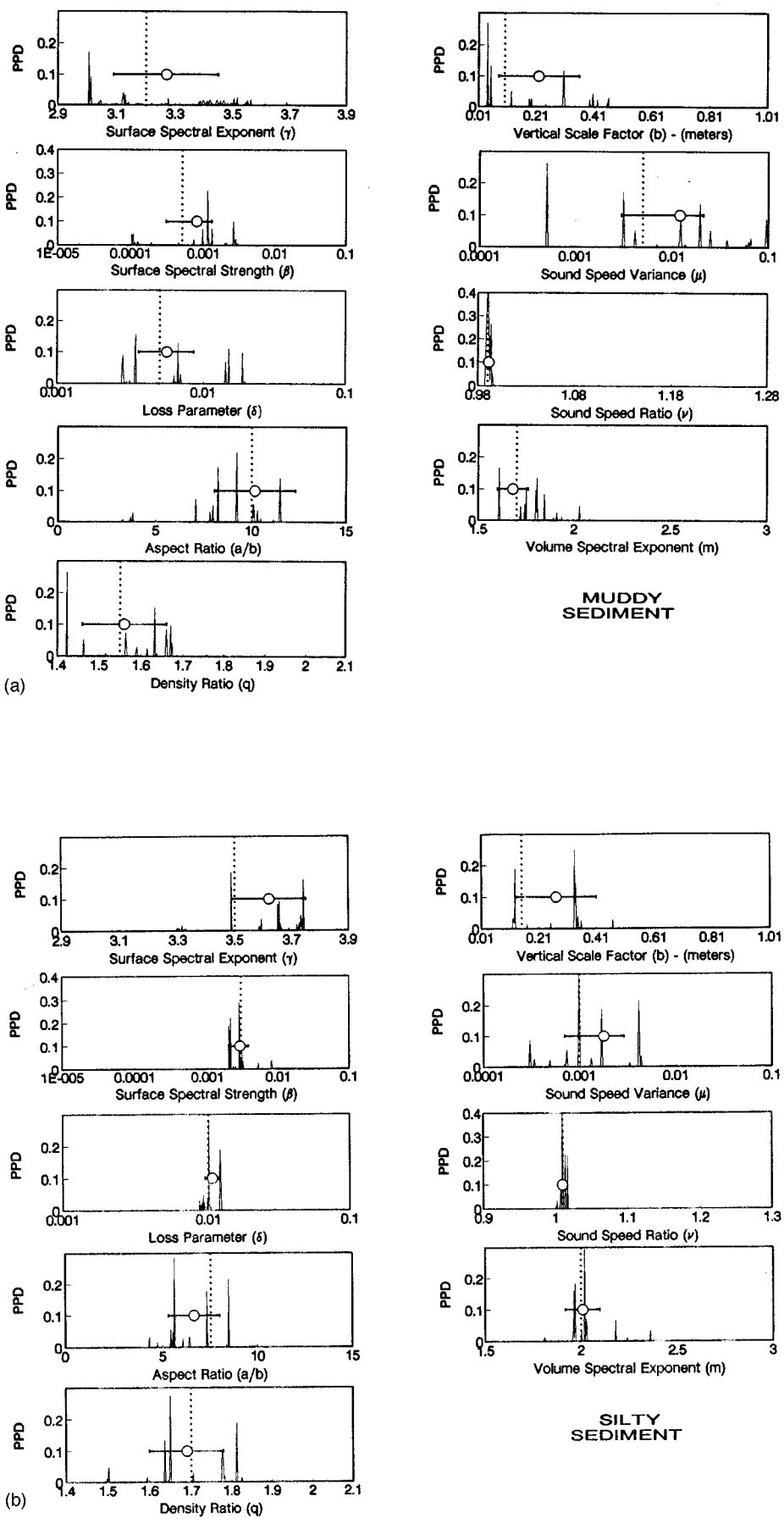


FIG. 5. *A posteriori* probability density functions of nine model parameters inverted for (a) muddy, (b) silty, and (c) sandy sediments. The GA-estimated mean and one standard deviation (horizontal solid line), and the true value (vertical dotted line) are also shown for each parameter.

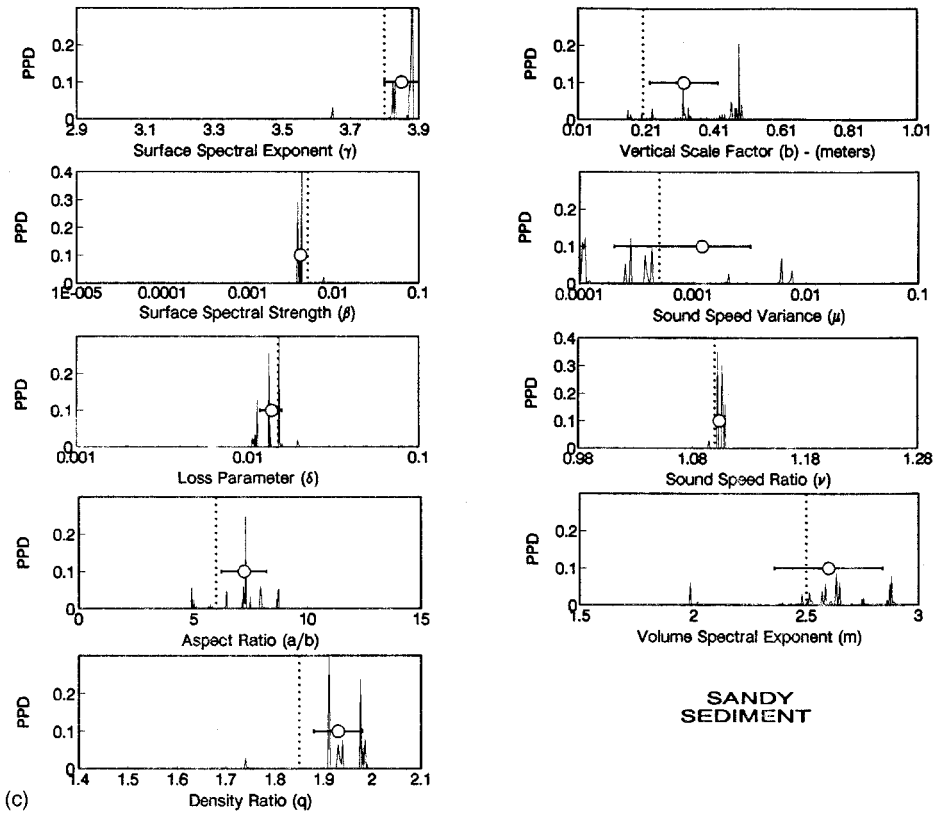


FIG. 5. (*Continued.*)

etc.). The resulting PPDs of eight model parameters (γ, β, δ, b, α, μ, ν, and m) for muddy, silty, and sandy sediments are shown in Fig. 6(a)–(c) respectively. The calculated mean and one standard deviation of the eight estimated parameters are also shown in Fig. 6(a)–(c). Again, in most cases the true values fall within one standard deviation of the estimated with generally smaller standard deviations than those of the nine-parameter inversion [see Fig. 5(a)–(c)].

Figure 7(a)–(c) shows the normalized *a posteriori* covariance matrices (normalized by $\langle \mathbf{m} \rangle \langle \mathbf{m} \rangle^{\dagger}$) of the eight-

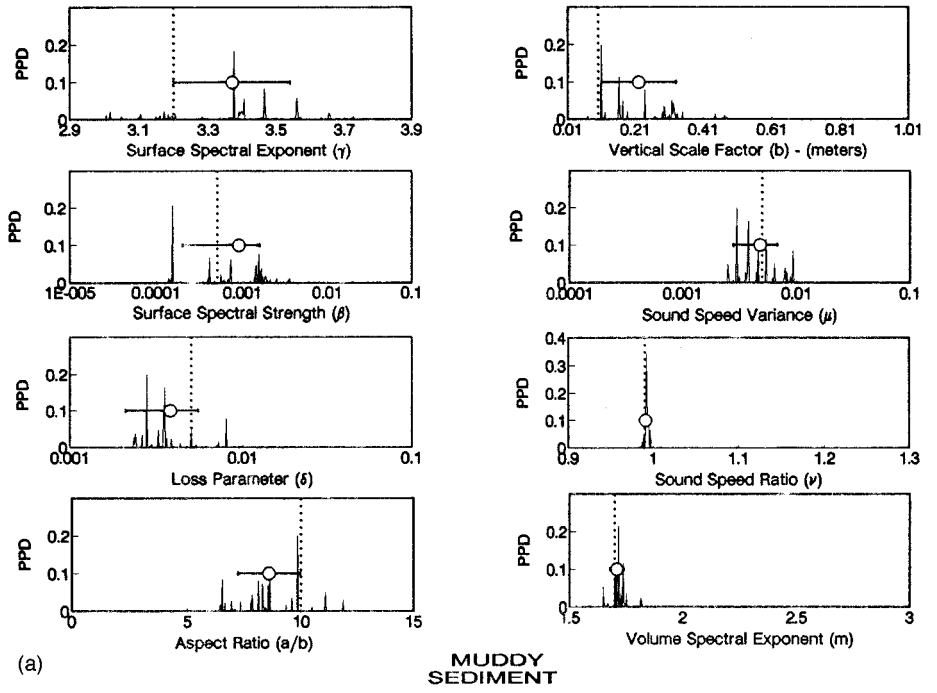


FIG. 6. *A posteriori* probability density functions of eight model parameters inverted for (a) muddy, (b) silty, and (c) sandy sediments. The GA-estimated mean and one standard deviation (horizontal solid line), and the true value (vertical dotted line) are also shown for each parameter.

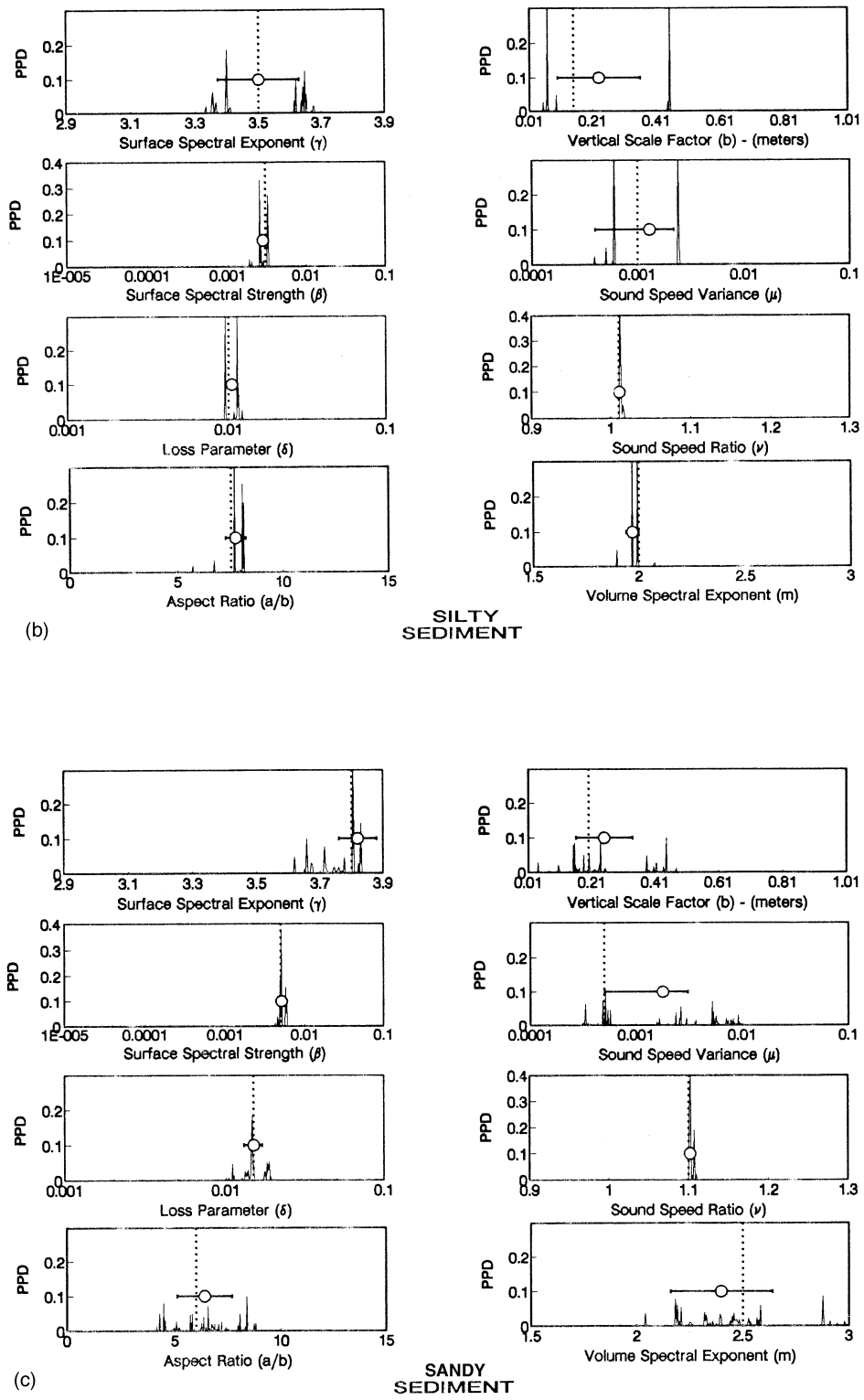


FIG. 6. (*Continued.*)

parameter inversion for three types of sediments. In Fig. 7(a), relatively large uncertainty of γ and β (as compared to those of silty and sandy sediments) indicates that scattering is less sensitive to interface parameters for muddy sediments. Similarly, relatively larger uncertainty of the volume parameters indicates scattering is less sensitive to volume parameters b and μ for sandy sediments [compare Fig. 7(c) with Fig. 7(a), (b)]. Figure 7(b) depicts an intermediate case in which scattering is sensitive to both volume and interface parameters for silty sediments (i.e., both interface and volume parameters contribute to the backscattering strength to some extent).

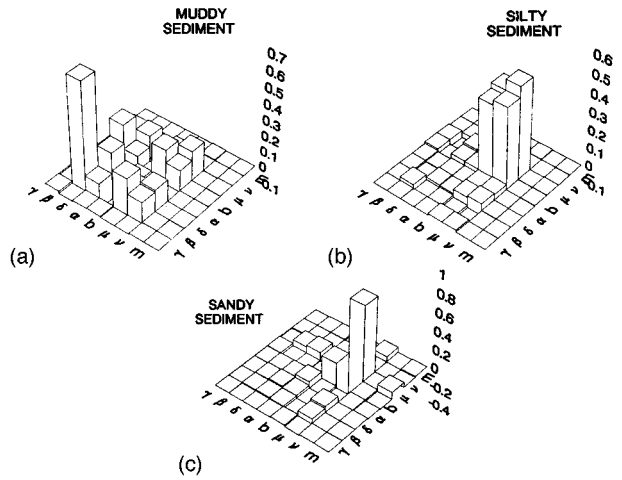


FIG. 7. *A posteriori* covariance matrices of eight model parameters calculated for (a) muddy, (b) silty, and (c) sandy sediments.

IV. EXPERIMENTAL RESULTS

A. Tomographic measurements of volumetric inhomogeneities in the sediment

Travel-time tomography was used to invert the structure of compressional wave speed between two vertical sediment probes. The results give an independent determination of volumetric parameters so that they can be compared with the backscattering inversion results. The tomographic inversion method utilizes the relation between observed travel-time T and sound-speed structure $v(\mathbf{r})$ through a line integral (Bregman *et al.*, 1989),

$$T=\int_{\Gamma}\frac{dl}{v(\mathbf{r})}. \tag{23}$$

Here l is the arc length along the ray and Γ is the raypath which is a function of sound-speed structure. Travel time is nonlinearly related to the reciprocal of the sound-speed structure since the raypath Γ also depends on the sound-speed structure. The problem can be solved linearly by applying some perturbations $\delta v(\mathbf{r})$ to an initial sound-speed structure $v_0(\mathbf{r})$ and fixed raypaths Γ_0. Then approximate differences in travel times are

$$\delta T=-\int_{\Gamma_0}\frac{\delta v(\mathbf{r})dl}{v_0^2(\mathbf{r})}, \tag{24}$$

where second- and higher-order perturbations are neglected. After parameterizing the model, Eq. (24) is expressed in matrix-vector notation as

$$\delta\mathbf{T}=\mathbf{A}\delta\mathbf{v}, \tag{25}$$

where $\delta\mathbf{T}$ are the travel-time differences between the data and those obtained by ray tracing, the matrix $\mathbf{A}$ has elements $a_{ij}=\partial T_i/\partial v_j$, and $\delta\mathbf{v}$ are the sound-speed perturbations which are related to travel-time differences through matrix $\mathbf{A}$. Matrix $\mathbf{A}$ is $m\times n$, where n is the number of grid points in the spatial domain and m is the number of measured travel-time data between sources and receivers. A damped least-squares (LSQR) method (Paige and Saunders, 1982) is used

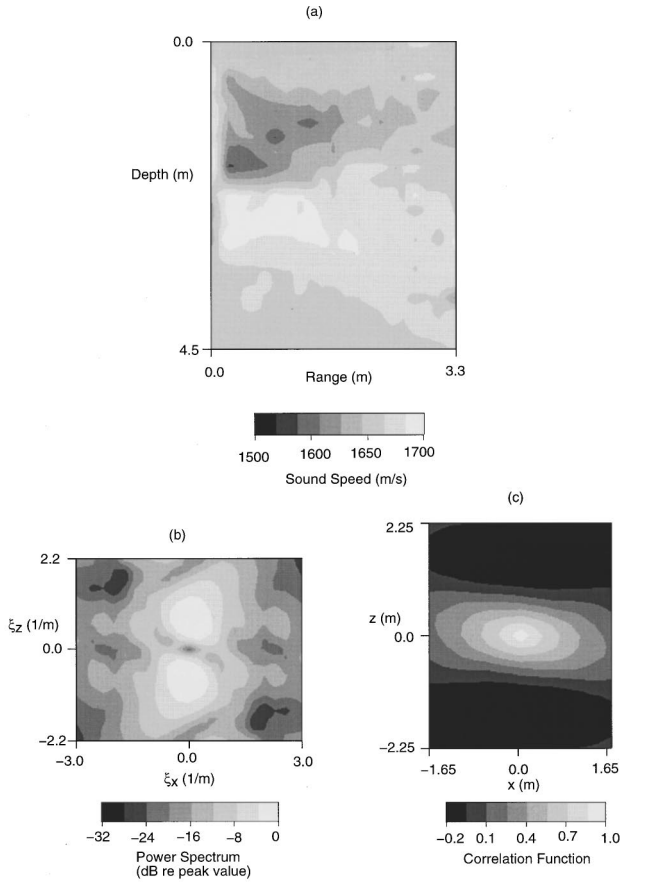


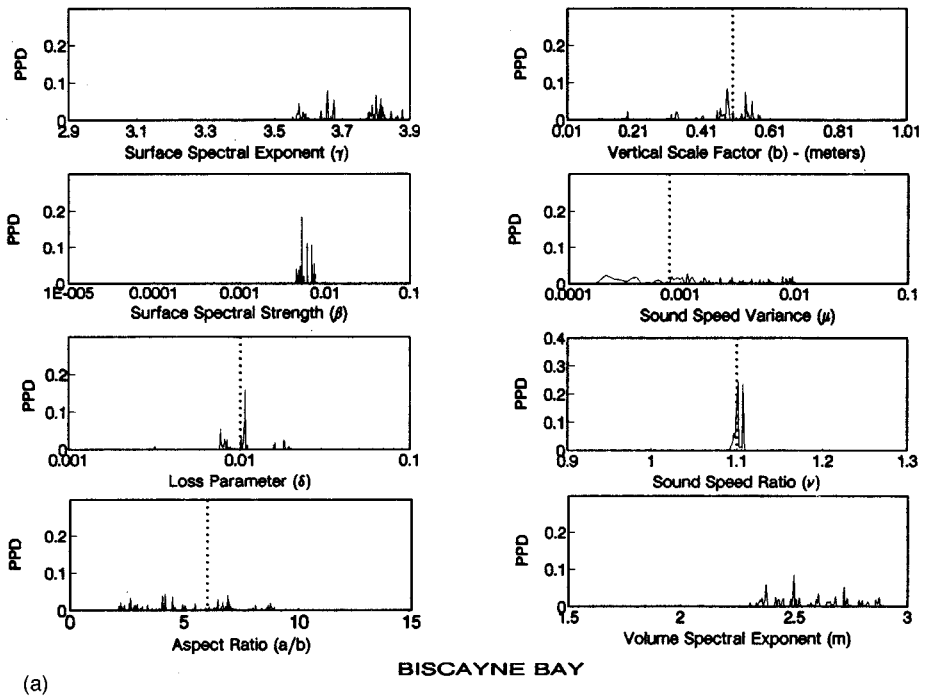
FIG. 8. (a) Tomographic inversion of 2-D sound-speed structure after the third and final iteration. (b) Wave number spectrum. (c) Correlation function of sound-speed inhomogeneities.

to calculate the sound-speed perturbation. The damped least-squares solution of the above equation is

$$\delta\mathbf{v}=(\mathbf{A}^T\mathbf{A}+\lambda^2\mathbf{I})^{-1}\mathbf{A}^T\delta\mathbf{T}, \tag{26}$$

where $\mathbf{I}$ is the identity matrix and λ is the damping parameter. The sound-speed distribution and corresponding raypaths are updated at each iteration, and the iteration process is repeated until the travel-time differences are less than the data error, or the change in sound-speed distribution is small. Details of the inversion and an explanation of the numerical code are given in Turgut (1993).

During the tomography experiment, multi-element sediment probes were buried as deep as 5 m in sandy sediments by using a hydraulic jet burial system (Turgut, 1993) with the assistance of SCUBA divers. An arbitrary function generator was interfaced with a personal computer to generate a frequency sweep or pseudo-random binary sequence code for the accurate measurement of travel times. Measured travel times and relative source–receiver positions were used in the tomographic inversion program. Figure 8(a) shows the inverted compressional sound-speed structure after the third and final iteration. It should be noted that the dimensions of the inverted cross section are 3.3 m×4.5 m and the spatial resolution of the tomographic inversion is about 0.15 m (determined by synthetic runs). The spatial resolution of the tomographic measurements is not sufficient to predict the



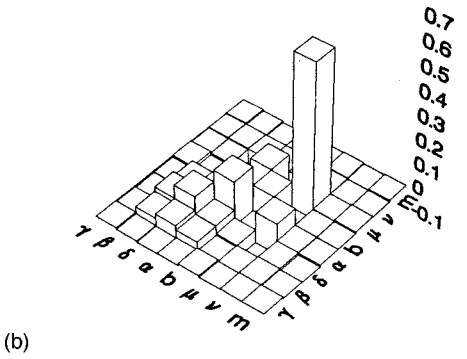


FIG. 9. (a) *A posteriori* probability density functions and (b) covariance matrices of eight model parameters inverted for Biscayne Bay sediment.

TABLE III. Comparison of GA-inverted model parameters and those obtained by tomographic measurements for Biscayne Bay sediments.

Parameters	Symbols	Biscayne Bay sediment	
		GA inversion	Tomography
Surface spectral exponent	γ	3.70±0.11	···
Surface spectral strength	β	0.0057±0.0009	···
Loss parameter	δ	0.0114±0.0036	0.01
Vertical scale factor	b	0.44±0.11	0.5
Aspect ratio	α	5.22±2.1	6.0
Sound-speed variance	μ	0.0011±0.0007	0.0008
Sound-speed ratio	ν	1.100±0.004	1.1
Volume spectral exponent	m	2.57±0.17	···

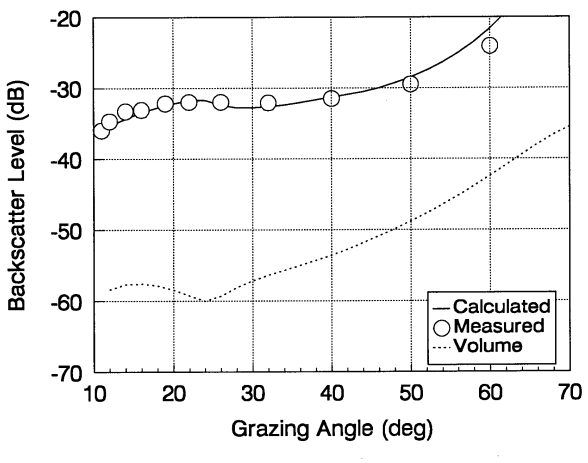


FIG. 10. Comparison of measured and model calculations of backscattering strength levels using the inverted mean model parameters for Biscayne Bay sediments (see Table III). Volume scattering contribution to the backscatter using the parameters obtained by tomographic measurements is also plotted.

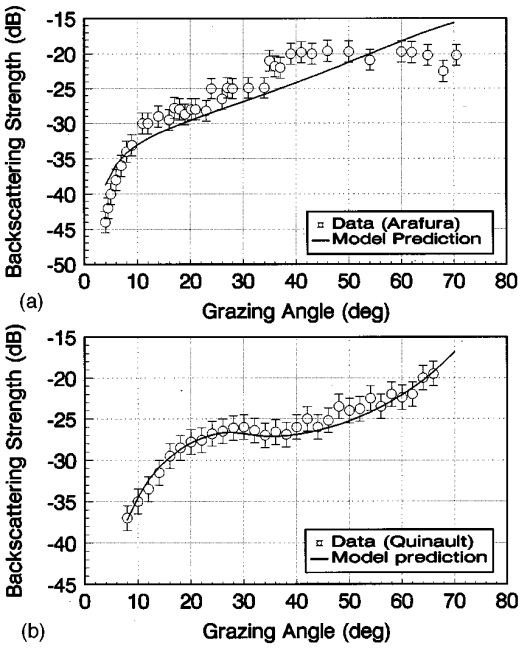
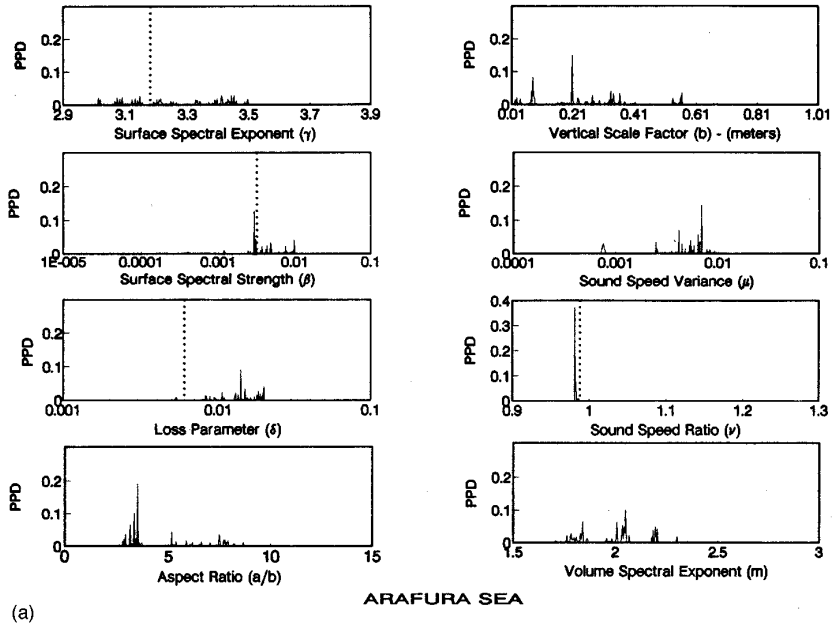


FIG. 11. Comparison of measured and model calculations of backscattering strength levels using the inverted mean model parameters (see Table IV) at (a) Arafura Sea and (b) Quinault Range sites.

parameters for volume scattering at frequencies higher than 6 kHz. In addition, at frequencies lower than 1 kHz, volume inhomogeneities deeper than 5 m might contribute to the scattering so that deeper tomographic images of sediments are needed. Also, as in our case, the distance between the probes should be comparable to the penetration depths of the probes so that a reasonable angular coverage of rays for the cross-well geometry is achieved. Larger separations cause poor horizontal resolution, which would artificially produce larger aspect ratios of sediment inhomogeneities. One also should keep in mind that the derivative matrix **A** is ill-conditioned and requires proper usage of damping parameter λ. Large damping values give a smoother solution which would produce small sound-speed variations. On the contrary, small damping values give an oscillatory solution which would produce large sound-speed variations. We found that the optimum damping parameter could be effectively obtained by inverting synthetic travel-time data for a synthetic (known) sound-speed structure having the real dimensions and positions of the sources and receivers.

Figure 8(b), (c) shows the 2-D wave number spectrum and correlation function of the inverted sound-speed inhomogeneities. The 2-D wave number spectrum is calculated by



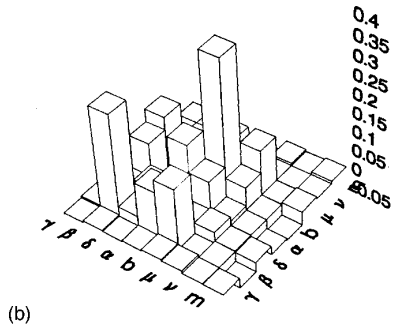


FIG. 12. (a) *A posteriori* probability density functions and (b) covariance matrices of eight model parameters inverted for Arafura Sea site.

TABLE IV. Comparison of GA-inverted model parameters and those by direct measurements (Jackson and Briggs, 1992) at Arafura Sea and Quinault Range sites.

		Arafura Sea		Quinault Range	
Parameters	Symbols	GA inversion	Measurements	GA inversion	Measurements
Surface spectral exponent	γ	3.26±0.14	3.18	3.57±0.06	3.67
Surface spectral strength	β	0.0056±0.0029	0.0032	0.0069±0.0004	0.0060
Loss parameter	δ	0.0153±0.0039	0.006	0.019±0.006	0.011
Vertical scale factor	b	0.24±0.13	⋯	0.32±0.17	⋯
Aspect ratio	α	4.85±2.1	⋯	5.85±1.6	⋯
Sound-speed variance	μ	0.0056±0.0017	⋯	0.011±0.0076	⋯
Sound-speed ratio	ν	0.981±0.006	0.988	1.134±0.005	1.113
Volume spectral exponent	m	1.97±0.15	⋯	2.32±0.26	⋯

taking the two-dimensional Fourier transform of the field after removal of a linear trend in the sound-speed structure. The linear trend represents a two-dimensional plane which is calculated by a least-squares curve-fitting method. Then, from the Wiener–Khinchin theorem, the correlation function is obtained by taking the inverse Fourier transform of the spectral density. Notice in Fig. 8(c) that a slight slope is present in the measured correlation function, representing a layer dipping (a few degrees) at the experimental site. This small layer dipping was negligible for backscattering calculations, however, it can be easily included in case of larger layer dips.

B. Inversion of Biscayne Bay backscattering data

An eight-element hydrophone array with 0.5-m spacing and an omnidirectional acoustic transducer were used for both monostatic backscattering measurements and spatial scattered field measurements. Half of the hydrophone array was buried in the sediment, an arrangement that enabled us to measure intrinsic attenuation. Intrinsic attenuation in the sediment was measured from the first-arrival pulses received by each buried hydrophone using the spectral ratio method (Turgut and Yamamoto, 1988). Only results from the monostatic scattering measurements are reported in this paper.

Backscattering strength was measured by using 5-kHz gated pulses at the same site where tomographic data of subbottom inhomogeneities were available. The scattering coefficient in this case is given as (Bunchuck and Zhitkovskii, 1980)

$$M_b = \frac{I_{\text{scat}}}{I_0} \frac{R^3}{\pi c_w \tau R_0^2}, \tag{27}$$

where R is the distance from the source to the insonified area at the water–sediment interface, c_w is the sound speed in water, I_0 is the intensity of the transmitted signal at a distance R_0 from the source, τ is the finite duration of the transmitted pulse, and I_{scat} is the intensity of the scattered signal. Since the transmitted and scattered signals were recorded by the same hydrophone, relative levels of I_{scat} and I_0 were used in the above equation. Then, backscattering strength was calculated using the relation $S_b = 10 \log_{10} M_b$.

Figure 9(a) shows the PPDs and normalized *a posteriori* covariance matrix of eight model parameters (γ, β, δ, b, α, μ, ν, and m) for the Biscayne Bay sediments. For comparison, results from the tomographic measurements (b, α, μ, and ν), and from the spectral ratio method (δ) are also plotted as vertical (dotted) lines. Geoacoustic parameters (δ and ν) are in good agreement with the tomographic measurements. However, although the tomographic measurements fall within one standard deviation of GA estimated volume parameters, standard deviations are rather large due to surface dominant scattering in sandy sediments. Figure 9(b) also shows the normalized *a posteriori* covariance matrix with larger uncertainties of volume parameters, which is a good indication of the importance of surface roughness at this site. In Table III, a comparison is made for the mean and one standard deviation of the inverted parameters with the tomographic measurements. Again, good agreement is shown between the tomographic measurements and the GA estimation of geoacoustic parameters (ν and δ). Although the standard deviations are large, the agreement for the volume parameters is still good especially considering the very small dependency of the total backscattering strength on these parameters in sandy sediments [see Fig. 2(c)]. Unfortunately, we did not have any quantitative measurements of interface parameters for the same site, so that GA estimation of surface parameters (γ and β) could not be compared.

The backscatter data and model prediction using the GA-estimated parameters are plotted in Fig. 10. The volume scattering contribution to the backscatter using the parameters obtained by tomographic measurements is also shown in Fig. 10. Since the calculated volume backscatter levels are much lower ($\approx$20 dB) than the measurements, the major source of scattering is attributed to interface roughness.

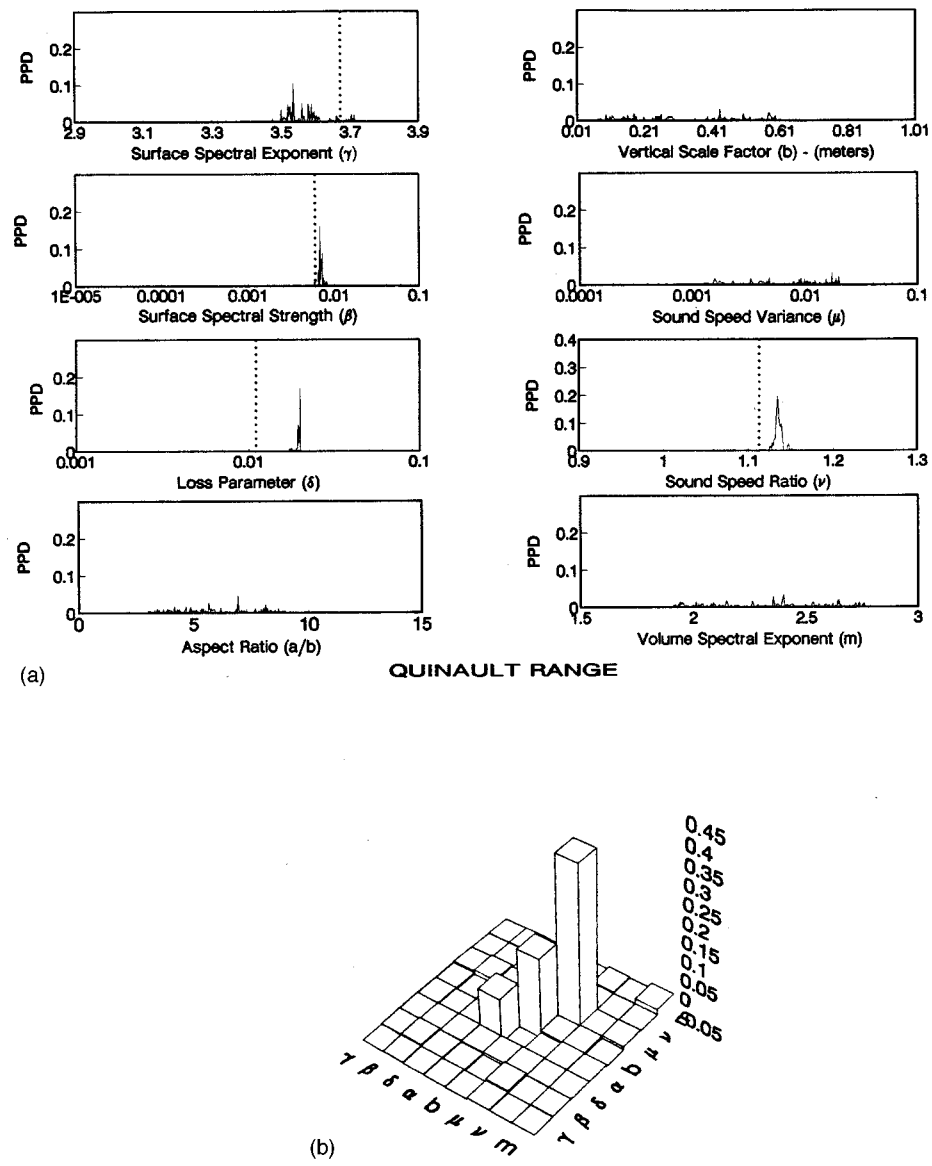


FIG. 13. (a) *A posteriori* probability density functions and (b) covariance matrices of eight model parameters inverted for Quinault Range site.

C. Inversion of Jackson–Briggs high-frequency backscattering data

Two high-frequency bottom backscattering data sets (Arafura Sea and Quinault Range) from Jackson and Briggs (1992) were also inverted, and the results were compared with their direct measurements (box-core and stereophotography) of geoacoustic and interface statistical parameters (see Table IV). The Arafura Sea site is characterized as silty/clayey sediment matrix with an isotropic bottom roughness. Their direct measurements at the Quinault Range indicated a silty sand bottom with slightly anisotropic interface roughness. However, the backscattering strengths measured in orthogonal directions were equal to within their experimental error. They attributed this discrepancy to the difficulties in estimating roughness spectra from the direct measurements especially considering a measured exponential strength value of 3.92 in the along-strike direction (acoustic propagation is parallel to the ripple crests). This corresponds to an unrealistic value of rms bottom slope of 33°. Since their backscattering data showed negligible anisotropy, they suggested using an exponential strength value of 3.67 (measured across-strike direction, in which propagation is perpendicular to the ripple crests) for both directions. Our GA-predicted value of 3.57 ± 0.07 agrees with their suggested value for the exponential strength. Other inverted geoacoustic and surface statistical parameters are comparable to their direct measurements representing typical characteristics of acoustically ''soft'' and ''hard'' sediments at the Arafura Sea and Quinault Range sites, respectively. In Fig. 11, Jackson–Briggs backscattering data and the model predictions using the inverted parameters are plotted for both Arafura Sea and Quinault site. The model predictions are usually within their specified 3-dB measurement errors. The calculated PPDs and normalized *a posteriori* covariance matrices for the Arafura Sea site are shown in Fig. 12(a),(b). Again, comparable uncertainties of the surface and volume parameters confirm the silty/clayey sediment type at this site. Figure 13(a),(b) shows the PPDs and normalized *a posteriori* covariance matrices

for the Quinault Range site. Reduced uncertainties of the surface parameters indicate that surface scattering is dominant at this site as expected for the relatively hard (sandy/silty) sediments.

V. CONCLUSIONS

It has been shown numerically that both interface and volume parameters of our extended scattering model can be estimated from the backscatter data using a parallel GA. Estimated mean models for three different types of sediments are generally very close to true models. The relative importance of each parameter, deduced from the diagonal elements of the covariance matrix, indicates the dominance of volume scattering for muddy sediments, and of interface scattering for sandy sediments. Statistical parameters inverted from the backscatter data collected in Biscayne Bay, Florida show the characteristics of a sandy bottom which was also confirmed by our laboratory analysis of sediment samples collected at the same site. Volume scattering parameters, measured by our sediment tomography method, also show the characteristics of a sandy bottom. Agreement between GA-estimated volume scattering parameters and tomographic measurements indicates that sediment inhomogeneities can be determined by carefully designed tomography and/or backscattering experiments. Similar simultaneous experiments conducted in muddy/silty sediments with small interface roughness would provide more insight to such comparison since the GA inverted volume scattering parameters should have less uncertainties for softer sediments. Inversion results of Jackson–Briggs high-frequency data also agree well with their directly measured geoacoustic and interface statistical parameter values. The frequency dependence of the volume backscattering strength can also be adequately modeled by choosing a proper values of the spectral exponent, the vertical scale factor, and the aspect ratio in our volume scattering model. Inversion of multi-frequency backscatter data can also be achieved with minor modifications in the algorithm so that the uncertainties in the estimation of both surface and volume parameters would be further reduced.

ACKNOWLEDGMENTS

The experimental part of this work was funded by the University of Miami Research Support Program while the author was at the University of Miami. Development of the bottom backscattering model and several improvements on the inversion algorithm were completed at the Naval Research Laboratory with the support by the Office of Naval Research. The author thanks Andrew Rogers and Steven Theophanis of the University of Miami for their help during the Biscayne Bay sediment tomography and backscattering experiment. Stephen Wolf's encouragement during the preparation of the manuscript is greatly appreciated.

APPENDIX

In this Appendix, the composite roughness model given by Mourad and Jackson (1988) and an extension of the volume scattering model given by Lyons *et al.* (1994) is summarized. Following the notation of Jackson *et al.* (1986), we assume a two-dimensional, isotropic surface roughness spectrum which obeys a power law for wave numbers comparable to the acoustic wave number:

$$W(\mathbf{K}) = \beta K^{-\gamma}, \tag{A1}$$

where $\mathbf{K}$ is the wave number vector of bottom relief features with the magnitude of $K = |\mathbf{K}|$, β is the spectral strength, and γ is the spectral exponent. In the composite roughness approximation, the small-scale roughness (Rayleigh–Rice) approximation is used with shadowing and large-scale bottom slope corrections (McDaniel and Gorman, 1983). The bottom backscattering cross section in the Rayleigh–Rice approximation is given as (Kuo, 1963)

$$\sigma_{\mathrm{pr}}(\theta) = 4k'^4 \sin^4\theta |Y(\theta)|^2 W(2k' \cos\theta), \tag{A2}$$

where k' is the acoustic wave number in the water. The complex function $Y(\theta)$ is defined as

$$Y(\theta) = \frac{(q-1)^2 \cos^2\theta + q^2 - \tau^2}{[q \sin\theta + (\tau^2 - \cos^2\theta)^{1/2}]^2}, \tag{A3}$$

where $q = \rho_0/\rho_f$ is the ratio of the density of the sediment to the density of the water, $\tau = (1 + i\delta)/\nu$ is the ratio of the complex wave number in the sediment to the real wave number in the water, δ is the loss parameter, and $\nu = c_0/c_f$ is the ratio of the sediment sound velocity to water sound velocity. The composite-roughness cross section is defined as

$$\sigma_{\mathrm{cr}}(\theta) = S(\theta, s) F(\theta, \sigma_{\mathrm{pr}}, s), \tag{A4}$$

where $S(\theta, s)$ is the shadowing correction, $F(\theta, \sigma_{\mathrm{pr}}, s)$ is a slope averaging integral, and s is the rms slope. A detailed evaluation of shadowing correction and slope averaging integral can be found in the APL-UW Technical Report-9407 (1994).

In the volume scattering model, the volume scattering cross section is calculated by using a modified version of the inhomogeneous continuum scattering model of Lyons *et al.* (1994) which is based on previous derivations by Chernov (1960), Tatarski (1961), and Morse and Ingard (1968). We assume the sediment is an inhomogeneous fluid with a small compressibility and density fluctuations around their average values:

$$\kappa = \kappa_0 + \delta\kappa, \quad \rho = \rho_0 + \delta\rho, \tag{A5}$$

with $|\delta\kappa/\kappa_0| \ll 1$, and $|\delta\rho/\rho_0| \ll 1$. Then, for harmonic acoustic motion with a single frequency, the modified Helmholtz equation for the pressure amplitude can be written as

$$\nabla^2 p + k^2 p = -k^2 \frac{\delta\kappa(\mathbf{r})}{\kappa_0} p - \nabla \cdot \left(\frac{\delta\rho(\mathbf{r})}{\rho_0} \nabla p \right), \tag{A6}$$

where k is the wave number in the sediment. As with most types of radiation problems, the above differential equation can be conveniently converted to an integral equation having components of volume and surface integrals. We use the free-space Green's function for an unbounded medium,

$$G(\mathbf{r}|\mathbf{r}_0) = \frac{e^{ikR}}{4\pi R}, \tag{A7}$$

where $R = |\mathbf{r} - \mathbf{r}_0|$.

In the scattering region, we assume the density and compressibility fluctuations are small, so the pressure distribution can be modeled as an incoming wave (Born approximation). In addition, a linear phase approximation, $|\mathbf{r}-\mathbf{r}_0| \approx |\mathbf{r}| - \hat{\mathbf{o}}\cdot\mathbf{r}_0$, can be made assuming the observation point is at the far field (Fraunhofer approximation). Then, for a plane-wave incident on a scattering volume, V, the scattered amplitude, $f(\hat{\mathbf{o}},\hat{\mathbf{i}})$ can be written as

$$f(\hat{\mathbf{o}},\hat{\mathbf{i}}) = \frac{k^2}{4\pi} \int_V \left[\frac{\delta\kappa(\mathbf{r})}{\kappa_0} + \frac{\delta\rho(\mathbf{r})}{\rho_0} \cos(\hat{\mathbf{o}},\hat{\mathbf{i}}) \right] \exp(i\mathbf{K}\cdot\mathbf{r}) dV, \tag{A8}$$

where $\mathbf{K} = k(\hat{\mathbf{i}} - \hat{\mathbf{o}})$ is the wave number transfer vector, and $\hat{\mathbf{i}}$ and $\hat{\mathbf{o}}$ are unit vectors in the directions of incoming and scattered waves, respectively.

Equation (A8) can be expressed in terms of the sound-speed variation in the medium as

$$f(\hat{\mathbf{o}},\hat{\mathbf{i}}) = \frac{k^2 V_{\kappa\rho}}{4\pi} \int_V \frac{\delta c(\mathbf{r})}{c_0} \exp(i\mathbf{K}\cdot\mathbf{r}) dV, \tag{A9}$$

where the parameter $V_{\kappa\rho}$ is defined as

$$V_{\kappa\rho} = \frac{-2[(\gamma_\kappa - 1) + (\gamma_\rho - 1)\cos(\hat{\mathbf{o}},\hat{\mathbf{i}})]}{(\gamma_\rho + \gamma_\kappa - 2)}, \tag{A10}$$

in which $\gamma_\kappa = \kappa_r/\kappa_0$ and $\gamma_\rho = \rho_r/\rho_0$.

The volume scattering cross section per unit volume is defined as

$$\Upsilon(\hat{\mathbf{o}},\hat{\mathbf{i}}) = (1/V)\langle |f(\hat{\mathbf{o}},\hat{\mathbf{i}})|^2 \rangle. \tag{A11}$$

Assuming the size of the scattering region is greater then the correlation distance, the scattering cross section per unit volume can be expressed by a power spectral density function, $S(\mathbf{K})$, as

$$\Upsilon = (\pi/2) k^4 V_{\kappa\rho}^2 S(\mathbf{K}), \tag{A12}$$

where

$$S(\mathbf{K}) = \frac{1}{8\pi^3} \int_V \left\langle \frac{\delta c(\mathbf{r})}{c_0} \frac{\delta c(\mathbf{r}+\mathbf{r}')}{c_0} \right\rangle \exp(i\mathbf{K}\cdot\mathbf{r}) dV. \tag{A13}$$

A general form of the von Karman correlation function is assumed to describe the isotropic sound-speed inhomogeneities in the sediment,

$$\left\langle \frac{\delta c(\mathbf{r})}{c_0} \frac{\delta c(\mathbf{r}+\mathbf{r}')}{c_0} \right\rangle = \frac{\mu}{2^{p-1}\Gamma(p)} \left(\frac{\mathbf{r}'}{r_0} \right)^p K_p \left(\frac{\mathbf{r}'}{r_0} \right), \tag{A14}$$

where $\mu = \langle [\delta c(\mathbf{r})/c_0][\delta c(\mathbf{r})/c_0] \rangle$ is the sound-speed variance, $K_p(\mathbf{r}'/r_0)$ is Bessel's function of the second kind of order p, and r_0 is the radial scale factor.

Considering a transversely isotropic structure of sediment having different scale factors in the horizontal and vertical directions, the resulting volume scattering cross section can be written as

$$\Upsilon = \frac{k^4 a^2 b V_{\kappa\rho}^2 \mu}{2\sqrt{\pi}} \frac{\Gamma(m)}{\Gamma(m-3/2)} [1 + 4k^2(a^2 \cos^2\theta_r + b^2 \sin^2\theta_r)]^{-m}, \quad m > 3/2, \tag{A15}$$

where b is the vertical scale factor, a is the horizontal scale factor, θ_r is the refracted angle in the sediment, and m is the spectral exponent ($m = p + 3/2$). After including refraction and transmission loss at the interface, volume scattering strength as a function of grazing angle is written as

$$\sigma_{\mathrm{pv}}(\theta) = \frac{\eta \Upsilon T(\theta)^2 \sin^2\theta}{2 \sin\theta_r}, \tag{A16}$$

where $\eta = [4\pi\Upsilon + \mathrm{Im}(k)]^{-1}$ is the inverse attenuation coefficient. Two-way transmissivity $T(\theta)$ in the above equation is defined as

$$T(\theta) = \frac{4q\nu \sin\theta \sin\theta_r}{(q\nu \sin\theta + \sin\theta_r)^2}. \tag{A17}$$

Shadowing and bottom slope corrections can be also incorporated in the calculation of volume scattering cross section analogous to Eq. (A4),

$$\sigma_v(\theta) = S(\theta,s) F(\theta, \sigma_{\mathrm{pv}}, s). \tag{A18}$$

APL-UW (**1994**). High-Frequency Ocean Environmental Acoustic Handbook, Technical Report APL-YW TR9407, AEAS 5501, October 1994.

Bourgeois, B., and Caruthers, J. (**1991**). ''Neural network parameter estimation for a bistatic scattering strength model,'' in *Oceans '91 Proceedings* (IEEE, New York), pp. 158–163.

Bregman, N. D., Bailey, R. C., and Chapman, C. H. (**1989**). ''Crosshole seismic tomography,'' Geophysics **54,** 200–215.

Briggs, K. B. (**1994**). ''High-frequency acoustic scattering from sediment interface roughness and volume inhomogeneities,'' NRL Report NRL/FR/7431-94-9617.

Bunchuck, A. V., and Zhitkovskii, Y. Y. (**1980**). ''Sound scattering by the ocean bottom in shallow-water regions,'' Sov. Phys. Acoust. **26,** 363.

Chernov, L. A. (**1960**). *Wave Propagation in a Random Medium* (McGraw-Hill, New York, translated from Russian by R. A. Silverman).

Crowther, P. A. (**1983**). ''Some statistics of the sea-bed and acoustic scattering there-from,'' in *Acoustics and the Seabed*, edited by N. G. Pace (Bath U. P., Bath, England).

De Jong, K. A. (**1975**). ''An analysis of the behavior of a class on genetic adaptive systems,'' Doctoral Dissertation, University of Michigan (University Microfilms No. 76-9381).

De Moustier, C., and Alexandrou, D. (**1991**). ''Angular dependence of 12-kHz seafloor acoustic backscatter,'' J. Acoust. Soc. Am. **94,** 522–531.

Gerstoft, P. (**1993**). ''Inversion of seismoacoustic data using genetic algorithms and *a posteriori* probability distributions,'' J. Acoust. Soc. Am. **95,** 770–782.

Goldberg, D. E., and Richardson, J. (**1987**). ''Genetic algorithms with sharing for multimodal function optimization,'' Proceedings of 2nd ICGA, 41–49.

Goldberg, D. E. (**1989**). *Genetic Algorithms in Search, Optimization, and Machine Learning* (Addison-Wesley, Reading, MA).

Gott, R. M., and Martinez, A. B. (**1993**). ''Estimation of composite roughness model parameters,'' in *Oceans '93 Proceedings* (IEEE, New York), pp. 444–449.

Jackson, D. R., and Briggs, K. B. (**1992**). ''High-frequency bottom backscattering: Roughness versus sediment volume scattering,'' J. Acoust. Soc. Am. **92,** 962–977.

Jackson, D. R., Winebrenner, D. P., and Ishimaru, A. (**1986**). ''Application of the composite roughness model to high-frequency bottom backscattering,'' J. Acoust. Soc. Am. **79,** 1410–1422.

Kuo, E. Y. (**1964**). ''Wave scattering and transmission at irregular surfaces,'' J. Acoust. Soc. Am. **36,** 2135.

Lyons, A. P., Anderson, A. L., and Dwan, F. S. (**1994**). ''Acoustic scattering from the seafloor: Modeling and data comparison,'' J. Acoust. Soc. Am. **95,** 2441–2451.

Mahfoud, S. W. (**1992**). ''Crowding and preselection revisited,'' in *Parallel Problem Solving from Nature*, edited by R. Manner and B. Manderick (Elsevier Science, Amsterdam, The Netherlands), Vol. 2, pp. 27–36.

Marsh, H. W. (**1961**). ''Exact solution of wave scattering by irregular surfaces,'' J. Acoust. Soc. Am. **33,** 330.

McDaniel, S. T., and Gorman, A. D. (**1983**). ''An examination of the composite roughness scattering model,'' J. Acoust. Soc. Am. **73,** 1476–1486.
Michalopoulou, Z. H., and Alexandrou, D. (**1996**). ''Bayesian modeling of acoustic signals for seafloor identification,'' J. Acoust. Soc. Am. **99,** 223–233.
Miller, B. L., and Shaw, M. J. (**1995**). ''Genetic algorithms with dynamic niche sharing for multimodal function optimization,'' IlliGAL Report No. 95010, University of Illinois at Urbana-Champaign, IL 61801.
Morse, P. M., and Ingard, K. U. (**1968**). *Theoretical Acoustics* (McGraw-Hill, New York).
Mourad, P. D., and Jackson, D. R. (**1989**). ''High-frequency sonar equation models for bottom backscatter and forward loss,'' in *Oceans '89 Proceedings* (IEEE, New York), pp. 1168–1175.
Paige, C. C., and Saunders, M. A. (**1982**). ''LSQR: An algorithm for sparse linear equations and sparse least squares,'' ACM Trans. Math. Software **8**, 43–71.
Sambridge, M., and Drijkoningen, G. (**1992**). ''Genetic algorithms in seismic waveform inversion,'' Geophys. J. Int. **109,** 323–342.
Sen, M. K., and Stoffa, P. L. (**1992**). ''Rapid sampling of model space using genetic algorithms: inversion of plane wave seismograms,'' Geophys. J. Int. **108,** 281–292.
Stoffa, P. L., and Sen, M. K. (**1991**). ''Nonlinear multiparameter optimization using genetic algorithms: Inversion of plane-wave seismograms,'' Geophysics **56,** 1794–1810.
Stockhausen, J. H. (**1963**). ''Scattering from the volume of an inhomogeneous half-space,'' Rep. No. 63/9, Naval Research Establishment, Canada.
Tang, D. (**1991**). ''Acoustic wave scattering from a random ocean bottom,'' Ph.D. Thesis, MIT-WHOI, Woods Hole, MA.
Tappert, F. (**1991**). *Lecture Notes*, March 13, University of Miami.
Tarantola, A. (**1987**). *Inverse Problem Theory: Methods for Data Fitting and Model Parameters Estimation* (Elsevier Science, Amsterdam, The Netherlands).
Tatarski, V. I. (**1961**). *Wave Propagation in a Turbulent Medium* (McGraw-Hill, New York).
Turgut, A. (**1991**). ''Measurements of scattering due to sea bottom roughness and sub-bottom sound speed fluctuations,'' J. Acoust. Soc. Am. **90,** 22 331.
Turgut, A. (**1993**). ''TOPOVEL: A FORTRAN Program for the tomographic inversion of sound speed and porosity structure of underground,'' Comput. Geosci. **14**, 511–531.
Turgut, A., and Yamamoto, T. (**1988**). ''Synthetic seismograms for marine sediments and determination of porosity and permeability,'' Geophysics **53,** 1056–1067.
Turgut, A., Yamamoto, T., Badiey, M., Trevorrow, M. V., and Conner, C. (**1988**). ''Bottom shear modulus profiler, a remote sensing instrument,'' Curr. Prac. New Technol. Ocean Eng. **13**, 87–95.
Urick, R. J. (**1983**). *Principles of Underwater Sound* (McGraw-Hill, New York).
von Karman, T. (**1948**). ''Progress in the statistical theory of turbulence,'' J. Mar. Res. **7**, 252–264.
Wood, A. B. (**1941**). *A Textbook of Sound* (Macmillan, New York).

Inverse scattering from arbitrary two-dimensional objects in stratified environments via a Green's operator

James W. Wiskin, David T. Borup, and Steven A. Johnson
Department of Bioengineering, University of Utah, Salt Lake City, Utah 84112

(Received 24 November 1996; accepted for publication 1 April 1997)

An important problem in geophysics, medical imaging, and nondestructive imaging today is the construction of a practical, accurate, and efficient means of imaging geophysical anomalies, tumours, or material defects in layered media. This paper discusses such a method. The use of a ''stratified Green's function'' for the solution of the forward problem is detailed. This forward problem is then incorporated into an efficient and accurate inversion algorithm based on optimization. The method is nonperturbative, unlike diffraction tomography, which relies on linearization to make the problem tractable. In the inversion, a pair of Lippmann–Schwinger-like integral equations are solved simultaneously via the Galerkin procedure for the unknown total internal fields and speed distribution. The computational burden is high, but made manageable by utilizing BiConjugate gradients, fast fourier transforms, and ''sinc'' basis functions to speed up the solution of the forward problem. The size and contrasts for which the method converges are substantially beyond the Born or Rytov approximations, and other methods heretofore reported in the literature. The convolutional character of the layered Green's function, and thus numerical efficiency, is preserved by careful construction based on known reflection coefficients. © *1997 Acoustical Society of America.* [S0001-4966(97)00708-X]

PACS numbers: 43.30.Ft, 43.30.Ma, 43.30.Pc, 43.30.Zk [JHM]

INTRODUCTION

One of the most important problems in marine geophysics, medical imaging, and nondestructive imaging today is the construction of a practical, accurate, and efficient means of imaging buried mines, geophysical anomalies, tumors, or material defects.[1–4] This paper will discuss a candidate for such a method, developed at the Center for Inverse Problems, Imaging, and Tomography (CIPIT) at the University of Utah. In particular, we detail the use of a ''stratified Green's function,'' G^L, for the solution of the forward problem, which is then incorporated into an inversion algorithm.

Inverse scattering algorithms may be broadly divided into two types: (1) those methods, such as the Marchenko integral equation,[5] which are exact; and (2) iterative, optimization-based procedures.[6–8] Our inversion is an iterative optimization procedure that requires the solution of many forward problems in its implementation. Due to the large number of forward problems that must be solved in order to image a scattering potential, the forward solution must be extremely fast. The work of Johnson *et al.*,[9] and in particular, the early work of Borup *et al.*,[10] has produced a very powerful technique for the fast solution of the forward problem and a concomitant iterative method for the imaging problem. See also Wilcox[11] for a clear summary of the inversion procedure in the case of homogeneous media (i.e., no stratification in the background medium).

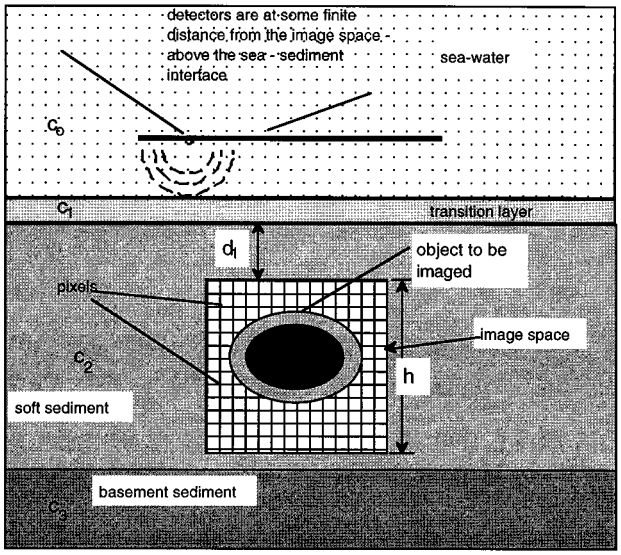
This paper deals exclusively with the scalar acoustic problem. The extension to the full elastic problem (i.e., including shear wave effects) will be carried out in a subsequent publication.

The problem solved in this paper is two-dimensional (2D), i.e., the object and field are assumed to have translational symmetry in the y-axis direction. Therefore, when we use coordinates x, y, and z in the discussion below, x and z describe the problem fully. The geometry relevant to both the inversion and the forward problem is displayed in Fig. 1(b), the image grid is nx by nz pixels in size, where nx is the number of pixels in the x (horizontal) direction, and nz is the number of pixels in the z (vertical) direction. A scatterer (e.g., a mine) is buried within the stratified medium, e.g., sediment, and the imaging grid is completely contained within a single layer, as shown in Fig. 1(b).

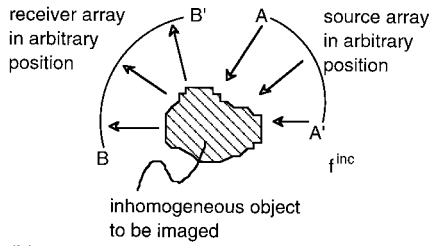
In both the forward and the inverse problem the acoustic wave speed, density, thickness, and attenuation of the layers in which the object is buried are known. The distinguishing characteristic of the inverse problem is this: the wave speed and attenuation of the inclusion, or mine (the object to be imaged) are unknown. The forward problem consists of determining the total field that results when a known incident field is projected on a scatterer with *known* parameter values (wave speed, etc.).

Thus this forward problem is analogous to the forward problem discussed in Ayme-Bellegarda,[12] which deals with the case of an object buried in a layered medium. References 6, 7, 8, and 12 serve as good introductions to this topic from a somewhat different point of view. In particular, consider the ''further work'' referred to in the conclusion of Ref. 13: ''Further work will include the investigation of an iterative scheme... to further enhance the reconstruction accuracy.'' In some sense these authors were anticipating the work carried out herein, in that the full nonlinear extension of the Born approximation is used.

In the following we utilize the 2-D acoustic stratified Green's functions (because of the abovementioned translation symmetry in the y direction). The fully 3-D case is discussed in detail elsewhere,[14] and in a future paper. The



(a)



(b)

FIG. 1. (a) The geometry for the forward and inverse problems. The sources and receivers are located along the thick solid bar shown above. The receivers are equally spaced at $\lambda/2$ spacing, where λ is the wavelength in the sediment at the highest frequency, e.g., $f_{\max}=6$ kHz implies $\lambda=0.247$ m. There are three sources equi-spaced along the array. (b) The source function giving the incident field is nonzero on AA′. The scattered field is measured at discrete locations spaced at $\lambda/2$ equal intervals along BB′. λ is the wavelength at the highest frequency.

fully elastic (i.e., including shear wave effects) equation is very similar to the acoustic case conceptually, but more difficult by virtue of its vector character, and the singular behavior of Green's kernel. It also will be addressed in a future publication. Recall from Ref. 15 that in the case of an inhomogeneous body residing within a homogeneous matrix, the governing equation is the Lippmann–Schwinger integral equation (Fredholm of the second kind):

$$f^{\text{inc}}_{\omega\theta}(\mathbf{r}) = f_{\omega\theta}(\mathbf{r}) - k_0^2 \int\int\int \gamma(\mathbf{r}) f_{\omega\theta}(\mathbf{r}) \times G_\omega(|\mathbf{r}-\mathbf{r}'|)\,dx'\,dy'\,dz', \tag{1}$$

with speed of sound $c(x,y,z)\equiv c(x,z)=\omega/k(x,y,z)$. $k_0^2=\omega^2/c_0^2$ is the wave number of the background (i.e., the layer containing the image grid) at frequency ω, where ω is the frequency of the interrogating wave; $k(x,y,z)$ is the spatially and frequency-dependent wave number of the body to be imaged; $f_{\omega\theta}(\mathbf{r})$ is the pressure (ultrasonic) field, with the spatial dependence on the vector $\mathbf{r}=(x,y,z)$ shown explicitly; $\gamma(\mathbf{r})$ is the object function, which will be reconstructed from the scattered data, and is given explicitly by $\gamma(x,z)\equiv(c_0^2/c(x,z)^2-1)$, and we have used the subscripts ω, θ to indicate the solution dependence on frequency (ω) and source position (θ). We assume that density is piecewise constant throughout each layer, and that the density of the object being imaged is the same as the layer in which it resides. In this case, and assuming $e^{i\omega t}$ time dependence, we have the following for a plane wave propagating in the $+x$ direction: $e^{i(\omega t-kx)}$. Defining the complex wave number as: $k\equiv k_R-i\alpha$ gives $e^{i(\omega t-kx)}=e^{i(\omega t-k_Rx)}e^{-\alpha x}$. This indicates that α is the amplitude attenuation coefficient in Np rs/m. The corresponding complex speed of sound is given by

$$c\equiv\frac{\omega}{k}=\frac{\omega}{k_R-i\alpha}=\frac{\omega(k_R+i\alpha)}{k_R^2+\alpha^2}.$$

Defining $c\equiv c_R+ic_I$, gives

$$c_R\equiv\omega\,\frac{k_R}{k_R^2+\alpha^2}=\frac{\omega}{k_R}\,\frac{1}{(1+\alpha^2/k_R^2)}$$

and

$$c_I=\omega\alpha\,\frac{1}{k_R^2+\alpha^2}=\frac{\omega}{k_R}\,\frac{\alpha}{k_R}\,\frac{1}{(1+\alpha^2/k_R^2)}.$$

These are exact expressions. Assuming that $\alpha/k_R\ll 1$, (a reasonable assumption) gives to second order: $c_R\approx(\omega/k_R)(1-\alpha^2/k_R^2)$, and $c_I\approx(\omega/k_R)(\alpha/k_R)$. This model of attenuation is acausal; however, it is a reasonable model for small α.

Symbolically, for a given speed of sound distribution: $\gamma(\mathbf{r})\equiv[c_0/c(\mathbf{r})]^2-1$, the linear operator $\phi_\omega(\gamma): f^{\text{inc}}_{\theta\omega}\rightarrow f_{\theta\omega}$ maps incident fields, $f^{\text{inc}}_{\theta\omega}$ at a particular frequency ω and direction of incidence, θ, to the corresponding total fields $f_{\theta\omega}$. This can be written: $[\phi_\omega(\gamma)]f^{\text{inc}}_{\theta\omega}=f_{\theta\omega}$. This operator will be given an explicit representation below. For now, the important point is that the inverse problem is given by the following: Given a ''range and domain restricted'' version of the forward map, $[\phi_\omega(\gamma)]$, determine the scattering parameters $\boldsymbol{\gamma}$. The range and domain restrictions arise from the following considerations. The sources and receivers are restricted to positions outside the object [see Fig. 1(a)]. In particular, the incident field is determined by the source functions on the arc AA′ in Fig. 1(a). The mapping $[\phi_\omega(\gamma)]$, then, can be considered as mapping that takes functions defined on the arc AA′ to functions defined on the arc BB′ (i.e., the scattered field). Furthermore the incident fields occur at a finite number of frequencies and arrive from a limited number of directions.

Several authors, such as Hofmann[16] and Natterer,[17] have utilized functional analysis to characterize the degree of ill-posedness of linear and nonlinear operators, and to give a clearer understanding of the relationship between the forward and the inverse problems. (Natterer[17] also includes a succint discussion of the relevant parts of Sobolev space formalism.) To this end, the forward operator can be expressed in terms of the following function spaces: (1) the function space of complex Lebesgue square integrable functions with compact support (e.g., BB′): $L_c^2\equiv L_c^2(\mathbf{R})$; (2) a subspace of $L_c^2\equiv L_c^2(\mathbf{R})$: functions with compact support, whose first and second derivatives are also square integrable, $H_c^2\equiv H_c^2(\mathbf{R})$; and (3) the function space of object functions $H_c^2\equiv H_c^2(\mathbf{R}^2)$—the subspace of functions in $H^2\equiv H^2(\mathbf{R}^2)$ which have compact support on $\mathbf{R}^2$.

The set of all linear maps from $L_c^2 \equiv L_c^2(\mathbf{R})$ to $H_c^2 \equiv H_c^2(\mathbf{R})$ is denoted by: $\mathrm{Hom}(L_c^2, H_c^2)$. In particular, an element of $\mathrm{Hom}(L_c^2, H_c^2)$ maps source functions on AA′ to scattered fields on BB′ in Fig. 1(a); that is, it solves the forward problem. In this notation the forward operator $\phi_\omega(\gamma) \in \mathrm{Hom}(L_c^2, H_c^2)$, and ϕ_ω itself is a map from speed of sound distributions $\gamma \in H_c^2$ to forward problem solvers:

$$\gamma \in H_c^2 \overset{\phi_\omega}{\rightarrow} [\phi_\omega(\gamma)] \in \mathrm{Hom}(L_c^2, H_c^2).$$

When the problem is discretized, the source functions $\in L_c^2$ and scattered fields $\in H_c^2$ have $O(n)$ degrees of freedom, so that $\mathrm{Hom}(L_c^2, H_c^2)$ has $O(n^2)$ degrees of freedom. The object functions γ have $O(n^2)$ degrees of freedom, where n is the edge dimension of the computational grid, $n \approx nx \approx nz$. Therefore the discretized version of ϕ_ω maps an $O(n^2)$ sized space to an $O(n^2)$ space.

In order to solve the inverse problem, we desire some approximation to $\phi_\omega^{-1}: \mathrm{Hom}(L_c^2, H_c^2) \rightarrow H_c^2$. This inversion cannot be carried out analytically or explicitly in the vast majority of cases; in fact, one rarely determines the operator ϕ_ω explicitly. Rather, by carefully analyzing the recorded scattered waveform for several different frequencies and source/receiver positions, the scattering parameters can be determined by a least-squares minimization. Before discussing the *full nonlinear* inversion formula, recall that the Born approximation can be used in the case of a homogeneous background when the scatterer is classified as being ''weak'' in the sense that the parameters of the scatterer differ only slightly from the background medium. A similar approximate solution can be obtained in the presence of a stratified medium. This solution is the approach taken in Ayme-Bellegarda.[11] By contrast, the algorithm outlined here will enable *quantitatively accurate* reconstructions of much higher contrast objects than is possible within the Born approximation (see Ref. 9 for further discussion).

The usefulness of this method is derived from the fact that computational ocean acoustics has reached a certain degree of maturity, as evidenced by the foundational book.[18] It is clear that different methods are applicable to varying scenarios, with different advantages and disadvantages arising in each cases. The inversion methods discussed here will provide a useful adjunct method in certain geometries (e.g., pure reflection mode) that preclude the use of other methods, such as the parabolic approximation.[19]

In Sec. I, the analogue of (1) for layered background media is derived; in Sec. II a suitable ''layered Green's function,'' $G_\omega(|\mathbf{r}-\mathbf{r}'|)$, is constructed. This Green's function can be considered as the point response at $\mathbf{r}$, for a continuous wave (cw) point source at $\mathbf{r}'$, of frequency ω. In Sec. III the ''layered media'' Lippmann–Schwinger equation is discretized, and some implementation issues are discussed. In Sec. IV a method of solution for the inverse problem is proposed and implementation issues are discussed. Some typical inversion scenarios are simulated in Sec. V. Section VI discusses some conclusions and further directions. The data for these examples were generated with the forward problem solution discussed in Ref. 20 which was verified in the case of a homogeneous background by using an analytic solution to generate the scattering from a homogeneous circular cylinder.[20]

It is important to note that the geometry shown in Fig. 1(b) has receivers and transmitters located on only *one* side of the image space. This increases the ill-posed nature of the problem dramatically. Results of simulations with sources and receivers surrounding the image space yield better results, but are not applicable to many situations, such as the scenarios presented in Sec. V.

I. DERIVATION OF THE STRATIFIED LIPPMANN-SCHWINGER EQUATION

The equation for the scattered field resulting from a point source located above the scattering object is

$$\nabla^2 f^{\mathrm{sc}}(x,y,z) + k_L^2(z) f^{\mathrm{sc}}(x,y,z) = -k_s^2(x,y,z) f(x,y,z), \quad (2a)$$

which can be rewritten as

$$\nabla^2 f^{\mathrm{sc}}(x,y,z) + k_L^2(z) f^{\mathrm{sc}}(x,y,z) = -\omega^2 \left(\frac{I}{c^2} - \frac{I}{c_L^2} \right) f(x,y,z), \quad (2b)$$

where the following definitions apply: The *total* wave number, $k(x,y,z)$ is given by $c^2(x,y,z) \equiv \omega^2 / k^2(x,y,z)$, where c is the spatially varying speed of sound [see Fig. 1(b)], and ω is the radial frequency $2\pi f$, with f being the frequency of the incident acoustic field. The total wave number $k(x,y,z)$ has been decomposed into $k_L(z)$, the wave number corresponding to the layers, and $k_s(x,y,z)$, the wave number in the arbitrary scattering object to be imaged in the following manner: $k^2(x,y,z) \equiv k_s^2(x,y,z) + k_L^2(z)$. Note that k_L is dependent only on the vertical direction z, whereas the wave number in the object k_s depends on all three coordinates, in general. The standard definition also holds for the background stratified medium speed $c_L(z)$: $c_L^2(z) \equiv \omega^2 / k_L^2(z)$. The speed of sound within the layer which hosts the image grid is c_{sc}. The stratified medium is assumed to have a constant speed of sound within each layer; in particular c_{sc} is constant [see Fig. 1(b)].

Equations (2a), (2b) can be rewritten as

$$\nabla^2 f^{\mathrm{sc}}(x,y,z) + k_L^2(z) f^{\mathrm{sc}}(x,y,z) = -(\omega^2 / c_{\mathrm{sc}}^2) \gamma_L f(x,y,z) \quad (3)$$

for an image grid located entirely within one layer. The ''stratified object function,'' γ_L, is defined by analogy with the free-space case discussed above: i.e., $\gamma_L(x,z) \equiv (c_{\mathrm{sc}}^2 / c^2(x,y,z) - 1)$.

The corresponding integral equation is then developed in the standard manner,[15] with the use of the Green's operator, G^L, which is defined as the solution to

$$\nabla^2 G^L(\mathbf{x};\xi) + k_L^2(z) G^L(\mathbf{x};\xi) = -\delta(x-\xi, z-\zeta), \quad (4)$$

where $\mathbf{x} \equiv (x,z)$ is the response point, and $\xi \equiv (\xi,\zeta)$ is the position of the continuous wave (cw) source. Using (4) in (3) gives a Fredholm equation of the second kind:

$$f^{\text{in}}(x,y,z)=f(\xi,\psi,\zeta)$$
$$-k_{\text{sc}}^2\int_V \gamma_L f(\xi,\psi,\zeta)G^L(x,y,z|\xi,\psi,\zeta)d\xi\, d\psi\, d\zeta, \tag{5}$$

which looks very much like the Lippmann–Schwinger equation for the homogeneous background case. In its present form, however, this equation does not display any convolutional character, which is critical for numerical efficiency. Clearly the integral will be a convolution in the horizontal (x-y) plane, since the background material is homogeneous in this direction. The vertical (z) direction, however, does not seem to enjoy this property, due to the arbitrary distribution of layers above and below the layer which hosts the inhomogeneity. It *is* possible, however, after some manipulation to rewrite the stratified Green's operator as the sum of a correlation term and a convolution term in the vertical (z) direction:

$$\mathbf{G}^L(\mathbf{r}-\mathbf{r}',\ z|z')=\mathbf{G}_V(\mathbf{r}-\mathbf{r}',\ z-z')+\mathbf{G}_R(\mathbf{r}-\mathbf{r}',\ z+z'). \tag{6}$$

A justification for this statement will be given in Sec. II. This observation is the key to the development of a fast inversion algorithm, since using a simple change of variables in the z direction, one can write the above correlational term as a convolutional operator, thereby allowing the use of fast Fourier transforms (FFTs). Anticipating this result, we define the total scattered field as the sum of a convolutional and a correlational contribution:

$$f^{\text{sc}}(x,y,z)\equiv f_V^{\text{sc}}(x,y,z)+f_R^{\text{sc}}(x,y,z), \tag{7}$$

where

$$f_V^{\text{sc}}(x,z)=k_{\text{sc}}^2\int_V G_V(x-\xi,z-\zeta)\,\gamma_L(\xi,\zeta)f(\xi,\zeta)d\xi\, d\zeta. \tag{8a}$$

Similarly, by making the substitution, $\zeta\equiv h-z'$, it follows that the correlational part is given by

$$f_R^{\text{sc}}(x,z)=k_{\text{sc}}^2\int_V G_R(x-\xi,h+z-z')(\gamma_L f\ )\times(\xi,h-z')d\xi\, dz'. \tag{8b}$$

This can be written as a convolution by defining the operator $\mathbf{Z}$ as coordinate reflection in the vertical (z) direction: i.e., $\mathbf{Z}\phi(\xi,\psi,\zeta)\equiv\phi(\xi,\psi,h-\zeta)$, for arbitrary functions ϕ, and G-hat by: $\hat{G}_R(x,z)\equiv G_R(x,h+z)$, so that $f_R^{\text{sc}}(x,z)=k_{\text{sc}}^2\hat{G}_R*\mathbf{Z}(\gamma_L f\ )$, where * denotes ''2-D convolution in x and z.''

The Lippmann-Schwinger equation in the stratified media case, then, reads,

$$f^{\text{inc}}=(\mathbf{I}-k_{\text{sc}}^2\mathbf{G_V}*\gamma_L-k_{\text{sc}}^2\hat{\mathbf{G}}_R*\mathbf{Z}\gamma_L)f, \tag{9}$$

where * denotes 2-D convolution, and this symbolic notation should be read from right to left: that is, the pointwise multiplication by γ_L is carried out *first*, followed by convolution with the Green's functions $\mathbf{G_V}$ and $\mathbf{G_R}$.

II. CONSTRUCTION OF THE STRATIFIED GREEN'S FUNCTION

The construction of the stratified Green's operator (the Green's operator associated with the unblemished stratified medium) is discussed in detail in Ref. 14. In particular, the problems associated with the wave number (or slowness) integration, and the solutions to this problem are discussed therein. Therefore the details concerning the construction and discretization of the Green's function, *per se*, will merely be outlined below, and then applications will be given. The 2-D Green's function (cylindrical harmonic waves from a line source) can be decomposed into plane waves in the following way[21] (analogous to the Weyl–Sommerfeld decomposition of a 3-D point source):

$$-\frac{i}{4}H_0^{(2)}(k_{\text{sc}}|\mathbf{r}-\mathbf{r}'|)=-\frac{i}{4\pi}\int_{-\infty}^{\infty}\frac{e^{-i(k_x(x-x')+\sqrt{k_{\text{sc}}^2-k_x^2}|z-z'|)}}{\sqrt{k_0^2-k_x^2}}\,dk. \tag{10}$$

In this formula $H_0^{(2)}(k_0|\mathbf{r}-\mathbf{r}'|)$ is the Hankel function of the second kind of (0th order). It is chosen because its asymptotic behavior is commensurate with an outgoing wave at large $|\mathbf{r}-\mathbf{r}'|$ when an $e^{i\omega t}$ time dependence is used. Also, k_x is the horizontal wave number, determined by the relation $\mathbf{k}\equiv(k_x,k_z)=k\hat{\mathbf{k}}$, so that

$$k_x=\frac{\omega}{c_{\text{sc}}}\sin\,\theta_{\text{sc}}\equiv k_{\text{sc}}\sin\,\theta_{\text{sc}}. \tag{11}$$

Here, c_{sc} denotes the wave speed in the layer containing the image grid and k_{sc} is the total wave number of the layer that contains the body to be imaged. Here, $k_z\equiv\sqrt{k_{\text{sc}}^2-k_x^2}$ is the vertical wave number, and the physically relevant sheet of the Riemann surface is chosen in such a manner that $\text{Im}\,\sqrt{k_{\text{sc}}^2-k_x^2}\leqslant 0$. The horizontal wave number k_x (and also horizontal slowness $u=k_x/\omega$) is useful for stratified media, since by Snell's law the horizontal wave number k_x, and therefore the horizontal slowness u, are preserved as the plane wave propagates through an arbitrary number of layers. Similarly, the reflected waves also have the same value of horizontal slowness.

It is possible to construct, therefore, a reflection coefficient; call it $R^-\equiv R^-(u)$, which represents the total reflected wave from *all layers* below the one which contains the image grid with horizontal slowness u. This coefficient includes all multiple reflections and reverberations within the stratified structure residing below layer ''sc.'' This R^- is constructed recursively (see Ref. 21 for details). A similar construction will give the reflection coefficient, denoted R^+, for all multiple reflections and reverberations resulting from all layers situated above the layer which houses the inhomogeneity (which we have labeled layer ''sc'' for convenience). In order to have a single reference position, R^+ and R^- are both given with respect to the top of the layer containing the inhomogeneity. That is, R^+ represents the reflection coefficient for an upgoing wave at the interface between layer sc and sc-1. The construction is carried out for

two cases: (1) when the receiver location is above the scattering location; and (2) when the receiver is located below the scatterer.

A. Case 1

The response point is located at some point above the scatterer. The plane-wave decomposition can be written[15] as

$$\frac{1}{4i\pi}\int_{-\infty}^{\infty}\frac{e^{-i\omega u(x-x')+i\omega b_{\rm sc}(z-z')}}{\sqrt{1/\beta_{\rm sc}^2-u^2}}\,du = \int_{-\infty}^{\infty}e^{-i\omega u(x-x')}S^u e^{i\omega b_{\rm sc}(z-z_{\rm sc})}\,du \tag{12}$$

with the upward source amplitude, S^u, defined as

$$S^u(u)\equiv\frac{e^{-i\omega b_{\rm sc}(z'-z_{\rm sc})}}{4i\pi b_{\rm sc}}. \tag{13}$$

In this formula, the vertical slowness for layer ''sc'' is

$$b_{\rm sc}=\sqrt{1/c_{\rm sc}^2-u^2}\quad\begin{matrix}b_{\rm sc}>0,\quad \gamma\in\mathbf{R},\\ \text{Im}\,b_{\rm sc}<0,\quad \gamma\in\mathbf{C},\end{matrix} \tag{14}$$

where $c_{\rm sc}$ is the wave speed in the layer containing the image grid. Integral (13) is interpreted as the upgoing wavefield due to a point source in 2-D (a line source in three dimensions) evaluated at some point above the source in the absence of layering. The total upward going component of the total wavefield at a point above the source when the layering is taken into account is given by

$$f^u(k_{\rm sc}|\mathbf{r}-\mathbf{r}'|)=\int_{-\infty}^{\infty}e^{i\omega u(x-x')}[1-R^-R^+]^{-1}\times[S^u+R^-S^d]e^{i\omega b_{\rm sc}(z-z_{\rm sc})}\,du. \tag{15}$$

Similarly, the total downward going component, including all reverberations from the layer interfaces, is given by the formula

$$f^d(k_{\rm sc}|\mathbf{r}-\mathbf{r}'|)=\int_{-\infty}^{\infty}e^{i\omega u(x-x')}[1-R^-R^+]^{-1}R^+\times[S^u+R^-S^d]e^{-i\omega b_{\rm sc}(z-z_{\rm sc})}\,du. \tag{16}$$

More care must be exercised in the fully elastic case. In this case mode conversion between compressional and vertically polarized shear waves (P-SV mode conversion) requires that R^+ and R^- are matrix valued functions of u, but similar representations are possible.

At this point, the definitions of S^u and S^d are substituted into integrals (15) and (16), the exponential terms are recombined, upgoing and downgoing field components are added, and the functions $C(u)\equiv(1-\mathbf{R}^-\mathbf{R}^+)^{-1}S^c$ and $S^c=i\omega u/b_{\rm sc}$ are defined. The resulting expression for the field is by definition, the sum of the convolutional and correlational parts of the layered Green's function:

$$G_V(x-x';|\Delta z-\Delta z'|)=\int_{-\infty}^{\infty}e^{i\omega u(x-x')}C(u)[e^{-i\omega b_{\rm sc}|\Delta z-\Delta z'|}+R^+R^-e^{i\omega b_{\rm sc}|\Delta z-\Delta z'|}]du, \tag{17a}$$

$$G_R(x-x';\Delta z+\Delta z')=\int_{-\infty}^{\infty}e^{i\omega u(x-x')}C(u)[R^+e^{-i\omega b_{\rm sc}(\Delta z+\Delta z')}+R^-e^{i\omega b_{\rm sc}(\Delta z+\Delta z')}]du. \tag{17b}$$

That is, the complete Green's function is given by the sum

$$G(x,z;x',z')\equiv G_{\mathbf{V}}(x-x',|z-z'|)+G_R(x-x',z+z'),$$

where $-h\leqslant x-x'\leqslant h$ and $2d_1\leqslant\Delta z+\Delta z'\leqslant 2d_1+2h$. Also d_1 is the distance from the top of the layer sc (the layer which contains the object to be imaged) to the top of the image space [see Fig. 1(b)]. Also, h is the height of the image space $h=nx\cdot\delta$. Using the fact that $C(u)[e^{-i\omega b_{\rm sc}|z|}+R^+R^-e^{i\omega b_{\rm sc}|z|}]$ is even, we are able to rewrite

$$G_V(x;z)=2\int_0^{\infty}\cos(\omega ux)C(u)\times[e^{-i\omega b_{\rm sc}|z|}+R^+R^-e^{i\omega b_{\rm sc}|z|}]du,$$

with a similar result for G_R.

B. Case 2

Here the receiver point is below the scatter point, and begins with a similar integral to (13). The intermediate steps lead to similar results by using slightly different integrals. By judicious use of the absolute value signs as shown, the final result does in fact turn out to be identical to Case 1.

III. THE STRATIFIED MEDIUM LIPPMANN–SCHWINGER EQUATION AND ITS DISCRETIZATION

A. Operator form of the Lippmann–Schwinger equation

The stratified media Lippmann–Schwinger equation for an imbedded scattering cylinder of arbitrary cross section can be written as

$$f^{\rm sc}\begin{pmatrix}x\\z\end{pmatrix}=k_0^2\int_{x'=0}^{x_0}\int_{z'=0}^{h}\left[\hat{G}_R\begin{pmatrix}x-x'\\z-z'\end{pmatrix}\hat{f}\hat{\gamma}_L\begin{pmatrix}x'\\z'\end{pmatrix}-G_V\begin{pmatrix}x-x'\\z-z'\end{pmatrix}f\gamma_L\begin{pmatrix}x'\\z'\end{pmatrix}\right]dz'\,dx', \tag{18}$$

where the ''caret'' affects the functions and the Green's kernel in the following manner:

$$\hat{f}\begin{pmatrix}x\\z\end{pmatrix}\equiv f\begin{pmatrix}x\\h-z\end{pmatrix},\quad \hat{\gamma}\begin{pmatrix}x\\z\end{pmatrix}\equiv\gamma\begin{pmatrix}x\\h-z\end{pmatrix},$$

and

$$\hat{G}_R\begin{pmatrix}x\\z\end{pmatrix}\equiv G_R\begin{pmatrix}x\\h+z\end{pmatrix}. \tag{19}$$

In symbolic notation, then, the Lippmann–Schwinger equation is

$$f^{\text{inc}} = f - k_{\text{sc}}^2 \mathbf{G}_V * (\gamma^L f) - k_{\text{sc}}^2 \mathbf{G}_R * \mathbf{Z}(\gamma^L f), \tag{20a}$$

or, provided that the equation is read from right to left (i.e., in operator theoretic format):

$$f^{\text{inc}} = (\mathbf{I} - k_{\text{sc}}^2 \mathbf{G}_V * \gamma^L - k_{\text{sc}}^2 \mathbf{G}_R * \mathbf{Z}\gamma^L) f. \tag{20b}$$

This is the operator equation that is solved by means of the biconjugate gradient (BCG) algorithm to give the total internal field. In (20) the stratified Green's operator $\mathbf{G}^L * \equiv \mathbf{G}_V * + \hat{\mathbf{G}}_R * \mathbf{Z}$ has been decomposed into its constituent parts and $\mathbf{Z}$ is the reflection operator in the z direction given by the identity

$$(\mathbf{Z}(\gamma^L f))(z) \equiv (\gamma^L f)(h-z) \equiv \gamma^L(h-z) \cdot f(h-z),$$

the $\mathbf{r} = (x,y)$ dependence being suppressed since it is unaffected by the $\mathbf{Z}$ operator. The convolutions * are carried out by means of fast Fourier transforms (FFTs). The actual inversion of the matrix operator in (20) is never explicitly carried out, since BCG is used to iteratively solve the system.

BCG is an efficient and reliable algorithm to use in the solution of the forward problem, i.e., the determination of the total field f from (20) above. It is substantially faster than conjugate gradients, and appears to be more reliable than the conjugate gradient squared (cgs) algorithm. Recent numerical experiments indicate that the ''stabilized biconjugate gradient method'' is an even more efficient algorithm for our forward problems, but the present examples were carried out with biconjugate gradients.

B. Closed form expressions for the discretized stratified Green's operator

Equation (20) can be discretized by the method of moments with combined ''sinc'' and tent basis functions and testing with the Dirac distribution (i.e., point matching). There are several reasons for using the ''sinc'' basis functions: (1) The use of ''sinc'' basis functions is an effective means of ameliorating the singular self-term in the Green's function. The Paley–Wiener theorem guarantees the smoothness of a spatial-frequency bandlimited function; therefore, by convolving with the ''sinc'' function, and thus multiplying by a rect function in the frequency domain (i.e., bandlimiting it), the singularity in the spatial domain must, perforce, be eliminated. (2) ''Sinc'' basis functions enjoy certain optimal properties with regard to discretization. The authoritative text on the use of ''sinc'' methods for approximation and other problems[22] discusses in detail the optimal approximation qualities of ''sinc'' functions. In particular, it allows a spatial sampling of $\lambda/2$, where λ is the wavelength at the highest frequency. (3) The cardinality property of ''sinc'' basis functions guarantees that the Galerkin method is equivalent to the method of moments with point matching, making its numerical implementation relatively straightforward. In actuality, ''sincs'' allow $\lambda/2$ sampling of a bandlimited function; however, the presence of the product $G(\mathbf{r},\mathbf{r}')\gamma(\mathbf{r}')f(\mathbf{r}')$ in the Lippmann–Schwinger-type integral argues heuristically for the $\lambda/4$ sampling.

Note that the forward (easier) part of the Paley–Wiener theorem says that since the ''sinc'' $[\text{sinc}(x) = \sin(\pi x)/\pi x]$ basis functions are *σ bandlimited*, any γ which is expanded in terms of them must be analytic and therefore cannot have compact support. Since our γ's are in fact limited in space, we are making a numerical approximation that is not, strictly speaking, mathematically valid. This approximation has been found surprisingly accurate, however, in numerical simulations.

Using the definitions $S_{(n',\delta)}(x') = \text{sinc}(x' - n'\delta)$, and the ''tent'' or ''chapeau'' function:

$$\Lambda_\delta(z) \equiv \begin{cases} 1 + \dfrac{z}{\delta}, & -\delta \leqslant z \leqslant 0, \\ 1 - \dfrac{z}{\delta}, & 0 \leqslant z \leqslant \delta, \end{cases} \tag{21}$$

and

$$\Lambda_{(m',\delta)}(z') = \Lambda_\delta(z' - m'\delta),$$

and the substitution:

$$\gamma_L\begin{pmatrix} x' \\ z' \end{pmatrix} f\begin{pmatrix} x' \\ z' \end{pmatrix} = \sum_{n',m'} S_{(n',\delta)}(x') \Lambda_{(m',\delta)} \times (z')\, \gamma_L\begin{pmatrix} n'\delta \\ m'\delta \end{pmatrix} f\begin{pmatrix} n'\delta \\ m'\delta \end{pmatrix},$$

the contribution to the scattered field from the convolutional part of the stratified Green's operator becomes:

$$\sum_{n',m'} \gamma_L\begin{pmatrix} n'\delta \\ m'\delta \end{pmatrix} f\begin{pmatrix} n'\delta \\ m'\delta \end{pmatrix} \int_{x'=0}^{x_0} \int_{z'=0}^{h} G_V\begin{pmatrix} x-x' \\ z-z' \end{pmatrix} \times S_{(n',\delta)}(x') \Lambda_{(m',\delta)}(z')\, dz'\, dx'.$$

A similar contribution occurs from the correlation term.

The piecewise linear ''tent'' functions are used in the vertical direction for the following reason. If the sinc basis functions were used in vertical and horizontal directions, these ''sinc'' basis functions would have to be integrated in both the vertical and horizontal directions against the stratified Green's function G^L. This integration can be carried out analytically in the horizontal direction; however, it cannot be carried out analytically in the vertical direction because of the arbitrary distribution of layer speeds above and below the layer which contains the object to be imaged. By using the ''tent'' or ''chapeau'' functions in the vertical direction, the vertical part of the integration can also be carried out analytically. It is possible, of course, to approximate the vertical integral in the case of the sinc functions numerically.

Upon discretizing x and z:, we have, then

$$f^{\text{inc}}\begin{pmatrix} n\delta \\ m\delta \end{pmatrix} \equiv f_{nm}^{\text{inc}} = f\begin{pmatrix} n\delta \\ m\delta \end{pmatrix} - k_{\text{sc}}^2 \int_{x'}^{h} \int_{z'=0}^{h} \gamma_L\begin{pmatrix} x' \\ z' \end{pmatrix} \times f\begin{pmatrix} x' \\ z' \end{pmatrix} G^L\left(\begin{matrix} n\delta \\ m\delta \end{matrix}\middle|\begin{matrix} x' \\ z' \end{matrix}\right) dx'\, dz', \tag{22a}$$

or, upon using the convolutional form of the Green's operator, as described above:

$$f^{\mathrm{sc}}\begin{pmatrix} n\delta \\ m\delta \end{pmatrix} = \int_{x'=0}^{x_0}\int_{z'=0}^{h}\Bigg[\hat{G}_R\begin{pmatrix} n\delta - x' \\ m\delta - z' \end{pmatrix}\hat{\gamma}_L\hat{f}\begin{pmatrix} x' \\ z' \end{pmatrix} + G_V\begin{pmatrix} n\delta - x' \\ m\delta - z' \end{pmatrix}\gamma_L f\begin{pmatrix} x' \\ z' \end{pmatrix}\Bigg]dz'\ dx'. \tag{22b}$$

Therefore, we are able to write

$$f^{\mathrm{sc}}\begin{pmatrix} n\delta \\ m\delta \end{pmatrix} = f_V^{\mathrm{sc}}\begin{pmatrix} n\delta \\ m\delta \end{pmatrix} + f_R^{\mathrm{sc}}\begin{pmatrix} n\delta \\ m\delta \end{pmatrix}$$

for

$$\begin{cases} n = 0,\ldots,N-1 \\ m = 0,\ldots,M-1 \end{cases}$$

where

$$f_R^{\mathrm{sc}}\begin{pmatrix} n\delta \\ m\delta \end{pmatrix} \equiv k_0^2 \sum_{n',m'} \hat{f}\hat{\gamma}_L\begin{pmatrix} n'\delta \\ m'\delta \end{pmatrix}\int_{x'=0}^{x_0}\int_{z'=0}^{h}\hat{G}_R \times\begin{pmatrix} (n-n')\delta - x' \\ (m-m')\delta - z' \end{pmatrix}\mathrm{sinc}(x')\Lambda_\delta(z')dz'\ dx', \tag{23}$$

and a similar result holds for the convolutional contribution. These integrals can be written as a discrete convolution when the following definition of the discretized stratified Green's function is used:

$$G^V_{\pi/\delta}\begin{pmatrix} n \\ m \end{pmatrix} \equiv k_{\mathrm{sc}}^2\int_{x'=-\infty}^{+\infty}\int_{z'=-\infty}^{+\infty} G_V\begin{pmatrix} n\delta - x' \\ m\delta - z' \end{pmatrix} \times \mathrm{sinc}(x')\Lambda_\delta(z')dz'\ dx', \tag{24}$$

and the discretized convolutional contribution to the discretized integral equation becomes:

$$f_V^{\mathrm{sc}}\begin{pmatrix} n\delta \\ m\delta \end{pmatrix} = \sum_{n',m'} \gamma_L f\begin{pmatrix} n'\delta \\ m'\delta \end{pmatrix} G^V_{\pi/\delta}\begin{pmatrix} (n-n') \\ (m-m') \end{pmatrix}, \tag{25}$$

where

$$G^V_{\pi/\delta}\begin{pmatrix} n \\ m \end{pmatrix} = \int_{-\infty}^{+\infty} C(u)\mathbf{I}_m(u)\mathbf{I}_n(u)du,$$

and the integrals I_m and I_n are given by

$$\mathbf{I}_m(u) \equiv \int_{z'=-\infty}^{+\infty}[e^{-i\omega b_{\mathrm{sc}}|m\delta - z'|} + R^+R^-e^{i\omega b_{\mathrm{sc}}|m\delta - z'|}]\Lambda_\delta(z')dz',$$

and

$$\mathbf{I}_n(u) \equiv \int_{x'=-\infty}^{+\infty} e^{iu\omega(n\delta - x')}\mathrm{sinc}(x')dx'.$$

The subscript π/δ refers to bandlimiting in the x direction alone, the basis function is in the vertical direction, and the tent function is not analytic (its not even C^1), so the Paley–Wiener theorem does not stipulate bandlimiting in the vertical direction. There is a similar contribution from the discretized *correlational* term of the stratified Green's function, and the total scattered field is finally given by

$$f^{\mathrm{sc}}\begin{pmatrix} n\delta \\ m\delta \end{pmatrix} = k_{\mathrm{sc}}^2\sum_{n',m'}\Bigg(\gamma_L f\begin{pmatrix} n'\delta \\ m'\delta \end{pmatrix} G^V_{\pi/\delta}\begin{pmatrix} (n-n') \\ (m-m') \end{pmatrix} + \hat{\gamma}_L\hat{f}\begin{pmatrix} n'\delta \\ m'\delta \end{pmatrix}\hat{G}^R_{\pi/\delta}\begin{pmatrix} (n-n') \\ (m-m') \end{pmatrix}\Bigg), \tag{26}$$

where

$$\hat{G}^R_{\pi/\delta}\begin{pmatrix} n \\ m \end{pmatrix} \equiv \int_{x'=0}^{x_0}\int_{z'=0}^{h}\hat{G}^R\begin{pmatrix} n\delta - x' \\ m\delta - z' \end{pmatrix}\mathrm{sinc}(x')\Lambda_\delta(z')dz'\ dx'.$$

In order to obtain closed forms for the discretized Green's operator, the integral (24) and the analogous integral for the correlation term, above, are evaluated to give: (1) the convolution term:

$$G^V_{\pi/\delta}\begin{pmatrix} n \\ m \end{pmatrix} \equiv G^V_1\begin{pmatrix} n \\ m \end{pmatrix} + G^V_2\begin{pmatrix} n \\ m \end{pmatrix},$$

where, for $m=0$:

$$\begin{cases} G^V_1\begin{pmatrix} n \\ m=0 \end{pmatrix} \\ G^V_2\begin{pmatrix} n \\ m=0 \end{pmatrix} \end{cases} = \frac{2}{\omega 4\pi i}\int_0^{u_\delta}\frac{\cos(u\omega\cdot n\delta)C(u)}{b_{\mathrm{sc}}} \times\begin{cases} 2R^+R^-\left(i\delta - \left(\dfrac{e^{i\omega b_{\mathrm{sc}}\delta}-1}{\omega b_{\mathrm{sc}}}\right)\right) \\ 2\left(-i\delta - \left(\dfrac{e^{-i\omega b_{\mathrm{sc}}\delta}-1}{\omega b_{\mathrm{sc}}}\right)\right) \end{cases} du, \tag{27}$$

and for $m\neq 0$:

$$\begin{cases} G^V_1\begin{pmatrix} n \\ m\neq 0 \end{pmatrix} \\ G^V_2\begin{pmatrix} n \\ m\neq 0 \end{pmatrix} \end{cases} = \frac{2}{\omega^2 4\pi i}\int_0^{u_\delta}\cos(u\omega\cdot n\delta)C(u) \times\frac{[2-e^{i\omega b_{\mathrm{sc}}\delta}-e^{-i\omega b_{\mathrm{sc}}\delta}]}{(b_{\mathrm{sc}})^2} \times\begin{Bmatrix} R^+R^-e^{i\omega b_{\mathrm{sc}}|m|\delta} \\ e^{-i\omega b_{\mathrm{sc}}|m|\delta} \end{Bmatrix} du: \tag{28a}$$

(2) the correlational term:

$$\hat{G}^R_{\pi/\delta}\begin{pmatrix} n \\ m \end{pmatrix} \equiv \hat{G}^R_1\begin{pmatrix} n \\ m \end{pmatrix} + \hat{G}^R_2\begin{pmatrix} n \\ m \end{pmatrix},$$

with the following definitions

$$\begin{cases} \hat{G}_1^R\begin{pmatrix} n \\ m \end{pmatrix} \\ \hat{G}_2^R\begin{pmatrix} n \\ m \end{pmatrix} \end{cases} = \frac{2}{\omega^2 4\pi i} \int_0^{u_\delta} \cos(u\omega n\delta) C(u)$$

$$\times \begin{cases} R^-(u) e^{i\omega b_{\rm sc}(2d_1+h+m\delta)} \\ R^+(u) e^{-i\omega b_{\rm sc}(2d_1+h+m\delta)} \end{cases}$$

$$\times \frac{[2-e^{-i\omega b_{\rm sc}\delta}-e^{i\omega b_{\rm sc}\delta}]}{(b_{\rm sc})^2}\, du, \tag{28b}$$

where d_1 is the distance from the top of the layer which contains the inhomogeneity to the top of the image space, and $-N_z+1 \leqslant m \leqslant N_z-1$ and $0 \leqslant n \leqslant N_x-1$, and the slowness u_δ is defined as: $u_\delta = 1/c_\delta = k_\delta(1/\omega) = (\pi/\delta)/\omega$ for wave number $k_\delta = \pi/\delta$, $C(u) \equiv (1-\mathbf{R}^-\mathbf{R}^+)^{-1} S^c$, where $S^c = i\omega u/b_{\rm sc}$, and $b_{\rm sc}=$ the vertical slowness for the shear wave velocity $\beta_{\rm sc}$, in the layer containing the scattering point.

The construction of the incident field from a point source located at a point above the layered medium would normally require separate calculation of a wave number (or slowness) integral. However, by placing the sources in the same position as the receivers, and using reciprocity of the Green's function, the incident field in the computational grid can be constructed directly from the scattering Green's function, $G^L \equiv G^V + G^R$, as the sum of a convolutional and correlational part. Furthermore, since it is the Fourier transform of the transformed correlational part of the Greens function, $\hat{G}^R \equiv \mathbf{Z}(G^R)$, as opposed to G^R itself, that is stored the following formula must be used for the actual computation of the incident wave emanating from a point source located at $(nx = n',\ nz = m')$:

$$f^{\rm inc}\begin{pmatrix} n \\ m \end{pmatrix} \equiv G^V\begin{pmatrix} n-n' \\ m-m' \end{pmatrix} + \hat{G}^R\begin{pmatrix} n-n' \\ m-((N_z-1)-m') \end{pmatrix} \quad \forall \begin{matrix} n=1,\ldots,N_x \\ m=1,\ldots,N_z \end{matrix}. \tag{29}$$

C. Distorted wave Born approximation and its relation to full inverse scattering

Physically, the distorted wave born approximation (DWBA) is similar to the Born approximation applied to the problem of finding the scattering object function γ. However, the incident field is replaced everywhere by the total field resulting from the layers alone (in the absence of the scattering object, i.e., $\gamma = 0$) but with layers present. In particular, the DWBA includes the interaction of the scattered field from the layering with the scattering potential γ; it does *not*, however, include the secondary scattering: the interaction of the part of the total field resulting from the presence of γ with the layering and multiple interactions after that. This is accomplished only through the stratified Green's operator, which may therefore be considered to be a completion of the DWBA. Numerical studies have shown that ignoring the multiple scattering effects yields a relatively poor *quantitative* image in the case of medium to high contrast (as expected).

IV. NONPERTURBATIVE IMAGING IN STRATIFIED MEDIA—MULTIPLE VIEW, MULTIPLE FREQUENCY CASE ALGORITHM

In the case that there are limited views available, due to experimental limitations, multiple frequencies must be used to increase the well posedness of the problem.

Writing (20) as

$$f = (\mathbf{I} - k_{\rm sc}^2 \mathbf{G}_V * \boldsymbol{\gamma}^L - k_{\rm sc}^2 \mathbf{G}_R * \mathbf{Z}\boldsymbol{\gamma}^L)^{-1} f^{\rm inc} \tag{30}$$

leads naturally to the definition of the ''forward map''

$$\boldsymbol{\Phi}_\theta(\boldsymbol{\gamma}) \equiv (\mathbf{I} - k_{\rm sc}^2 \mathbf{G}_V * \boldsymbol{\gamma}^L - k_{\rm sc}^2 \mathbf{G}_R * \mathbf{Z}\boldsymbol{\gamma}^L)^{-1} f_\theta^{\rm inc}, \tag{31}$$

where the subscript θ is used to indicate the incident field from direction θ. A further subscript, ω, serves to enumerate the different frequencies used in the inversion. The map $\boldsymbol{\phi}$ is holomorphic (in $\boldsymbol{\gamma}$), a fact which is used to reduce the computational effort by adapting the Hestenes–Saunders least-squares conjugate gradient method[23] to a holomorphic map.

To solve the multiple view and multiple frequency problem requires the solution of the system

$$\mathbf{r}_{\theta\omega}^{(n)} \equiv \boldsymbol{\phi}_{\theta\omega}(\boldsymbol{\gamma}^{(n)}) - \mathbf{d}_{\theta\omega} = 0 \quad \forall \begin{matrix} \theta=1,\ldots,\Theta \\ \omega=1,\ldots,\Omega \end{matrix}, \tag{32}$$

where, as stated, θ refers to the multiple views and the ω to the multiple frequencies available. The Newton–Raphson procedure is applied to this problem with the knowledge that applying a Gauss Newton algorithm to the associated least-squares problem would give the same result by virtue of the holomorphic nature of $\boldsymbol{\phi}$. This leads to the overdetermined system

$$\left(\frac{\partial \boldsymbol{\phi}_{\theta\omega}}{\partial \boldsymbol{\gamma}}\right) \cdot \delta\boldsymbol{\gamma}^{(n)} \approx -\mathbf{r}_{\theta\omega}^{(n)} \quad \forall \begin{matrix} \theta=1,\ldots,\Theta \\ \omega=1,\ldots,\Omega \end{matrix}, \tag{33}$$

which must be solved for the $\boldsymbol{\gamma}$-update $\delta\boldsymbol{\gamma}$. The *complex analytic* version of the Hestenes overdetermined conjugate gradient algorithm, adapted for least-squares problems, is used to iteratively solve this system. The formula for the Jacobian in the stratified medium situation in the presence of multiple frequencies is easily calculated:

$$\left(\frac{\partial \boldsymbol{\phi}_{\theta\omega}}{\partial \boldsymbol{\gamma}}\right) = \mathbf{T} \circ ((\mathbf{I} - \mathbf{G}_\omega \cdot [\boldsymbol{\gamma}])^{-1} (\mathbf{G}_\omega \cdot [\mathbf{f}_{\theta\omega}])) \quad \forall \begin{matrix} \theta-1,\ldots,\Theta \\ \omega=1,\ldots,\Omega \end{matrix}, \tag{34}$$

where $\mathbf{G}_\omega$ is now the *stratified* Green's function for the frequency ω (the L superscript on the stratified Green's function is suppressed in the above and that which follows).

Therefore, in effect, to determine the $\boldsymbol{\gamma}$-update $\delta\boldsymbol{\gamma}$, the multiple view problem for each particular frequency, that is, the overdetermined system is solved:

$$[\mathbf{T} \circ ([\mathbf{I} - \mathbf{G}_k \cdot [\boldsymbol{\gamma}]]^{-1} \mathbf{G}_k) \otimes \mathbf{I}_{\Theta\times\Theta}] \begin{bmatrix} [f_{1k}] \\ \vdots \\ [\mathbf{f}_{\Theta k}] \end{bmatrix} \delta\boldsymbol{\gamma}^{(n)} = -\begin{bmatrix} [\mathbf{r}_{1k}^{(n)}] \\ [\mathbf{r}_{\Theta k}^{(n)}] \end{bmatrix} \quad \forall k = 1,\ldots,\Omega. \tag{35}$$

Writing the overdetermined system in the following matrix notation (recall that $[\boldsymbol{\gamma}]$ is the diagonal matrix formed from the vector $\boldsymbol{\gamma}$, and all of the juxtapositions in the following formula are matrix multiplications):

$$\mathbf{T}\begin{bmatrix} ([\mathbf{I}-\mathbf{G}_1\cdot[\boldsymbol{\gamma}]]^{-1}\mathbf{G}_1)\otimes\mathbf{I}_{\Theta\times\Theta} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & ([\mathbf{I}-\mathbf{G}_\Omega\cdot[\boldsymbol{\gamma}]]^{-1}\mathbf{G}_\Omega)\otimes\mathbf{I}_{\Theta\times\Theta} \end{bmatrix}\begin{bmatrix} \begin{bmatrix}[\mathbf{f}_{11}]\\ \vdots \\ [\mathbf{f}_{\Theta 1}]\end{bmatrix} \\ \ddots \\ \ddots \\ \begin{bmatrix}[\mathbf{f}_{1\Omega}]\\ \vdots \\ \mathbf{f}_{\Theta\Omega}\end{bmatrix}\end{bmatrix}\delta\boldsymbol{\gamma}^{(n)}=-\begin{bmatrix} \begin{bmatrix}[\mathbf{r}_{11}^{(n)}]\\ \vdots \\ [\mathbf{r}_{\Theta 1}^{(n)}]\end{bmatrix} \\ \ddots \\ \ddots \\ \begin{bmatrix}[\mathbf{r}_{1\Omega}^{(n)}]\\ \vdots \\ [\mathbf{r}_{\Theta\Omega}^{(n)}]\end{bmatrix}\end{bmatrix}. \tag{36}$$

For multiple view and multiple frequency inversion, then, the inversion scheme reads:

(1) Choose an initial guess for the scattering potential, $\boldsymbol{\gamma}^{(0)}$. Set $n=0$

(2) Begin: solve the forward problems

$$(\mathbf{I}-\mathbf{G}[\boldsymbol{\gamma}])\mathbf{f}_{\theta\omega}=\mathbf{f}_{\theta\omega}^{\text{inc}}, \quad \forall\begin{matrix}\theta=1,\ldots,\Theta\\ \omega=1,\ldots,\Omega,\end{matrix}$$

using biconjugate gradients, and use the result to compute the scattered field $\boldsymbol{\phi}_{\theta\omega}(\boldsymbol{\gamma}^{(n)})$ and residual

$$\mathbf{r}_{\theta\omega}^{(n)}\equiv\boldsymbol{\phi}_{\theta\omega}(\boldsymbol{\gamma}^{(n)})-\mathbf{d}_{\theta\omega} \quad \forall\begin{matrix}\theta=1,\ldots,\Theta\\ \omega=1,\ldots,\Omega.\end{matrix}$$

(3) Determine the $\mathbf{L}_2$ norm of the total residual vector:

$$\mathbf{tr}^{(\mathbf{n})}\equiv\begin{bmatrix} \begin{bmatrix}[\mathbf{r}_{11}^{(n)}]\\ \vdots \\ [\mathbf{r}_{\Theta 1}^{(n)}]\end{bmatrix} \\ \ddots \\ \ddots \\ \begin{bmatrix}[\mathbf{r}_{1\Omega}^{(n)}]\\ \vdots \\ [\mathbf{r}_{\Theta\Omega}^{(n)}]\end{bmatrix}\end{bmatrix}. \tag{37}$$

(4) If $\|\mathbf{tr}^{(\mathbf{n})}\|^2\equiv\Sigma_{\omega=1}^{\Omega}\Sigma_{\theta=1}^{\Theta}\|\mathbf{r}_{\theta\omega}^{(n)}\|^2<\epsilon$ then stop.

(5) Solve the $N_d\Theta\Omega$ by N overdetermined system (36) in the least-squares sense for $\delta\boldsymbol{\gamma}^{(n)}$, using the conjugate gradient algorithm adapted for least-squares problems.

(6) Update $\boldsymbol{\gamma}$ via the formula

$$\boldsymbol{\gamma}^{(n+1)}=\boldsymbol{\gamma}^{(n)}+\delta\boldsymbol{\gamma}^{(n)}. \tag{38}$$

(7) Set $n=n+1$; go to begin.

The algorithm requires overdetermination, i.e., that: $nx\cdot nz\leqslant(2\cdot(nx+nz)-4)\cdot\Theta\Omega$. The actual values of the angles of incidence θ are chosen to diminish as much as possible the ill-conditioning of the problem. Equally spaced values $0\leqslant\theta\leqslant 2\pi$ are ideal, but experimental constraints may prohibit such values.

The adjoint is required in the computation of the Hestenes–Saunders conjugate gradients least-squares algorithm, and can easily be computed (with respect to the standard inner product on $\mathbf{C}^{\mathbf{N}}$).

V. RESULTS OF SIMULATIONS

A generic reflection-only geometry, showing the sources and receivers above the layers of sediment, is shown in Fig. 1(b). Figure 2(a) and (b) shows the gray scale of an original 2-D model, and the reconstruction using the FFT–BCG layered Green's function algorithm. There were 64 frequencies used in the simulation shown in Fig. 2, which were equally spaced from 0 to $f_{\max}=80$ kHz. Twelve sources were placed along the top of this image space, adjacent to it. The image space was surrounded by layers representing sea water and sediment with $c_0=1480$ m/s (sea water), $c_1=1500$ m/s (topmost sediment layer), $c_2=1680$ m/s (second sediment layer), $c_3=2500$ m/s [bottom sediment layer; see Fig. 1(b)]. The spatial sampling at the highest frequency was $\lambda/4$ and there were a total of 64 equally spaced frequencies used, with $f_{\min}=f_{\max}/64$. The original computer model is shown in Fig. 2(a) and its reconstruction is shown in Fig. 2(b). The lightest gray scale represents the highest speed of sound, with $\gamma=-0.2$, and the darkest shade has a $\gamma=0.2$. The gray scale reconstruction shown in Fig. 2(b) does not adequately represent the quantitative accuracy of the reconstruction, so cross sections at various depths are shown in Fig. 3(a)–(c). The image space itself is 128 by 32 pixels at $\lambda_{\max}/4$ sampling in size; the cross sections shown in Fig. 3(a)–(c) were through $nz=14$, 18, and 22, where nz is the pixel number in the vertical direction, where down is the positive direction.

The image space shown in Fig. 4 was sandwiched between layers with $c_1=1480$ m/s (sea water), and $c_2=1650$ m/s (sediment) to simulate the reconstruction of a buried mine simultaneously with a sloping sediment surface. It is considerably smaller than the previous image space, being only 64 by 16 pixels (16 by 4λ), but it does display the ability of the reconstruction algorithm to quantitatively reconstruct an uneven sedimentary interface. The light gray represents sea water, and the darker gray sediment. A buried mine is shown in cross section with $\gamma=-0.4$ representing a hard shell, and the interior object function γ has $\gamma=0.4$. Figure 5 is a cross section of the reconstruction shown in Fig. 4, which displays the quantitative accuracy of the reconstruction.

The question naturally arises as to whether the minima of the functional that are achieved are in fact global minima. Several options exist to determine this, but for practical problems it seems best to obtain a reasonable first guess us-

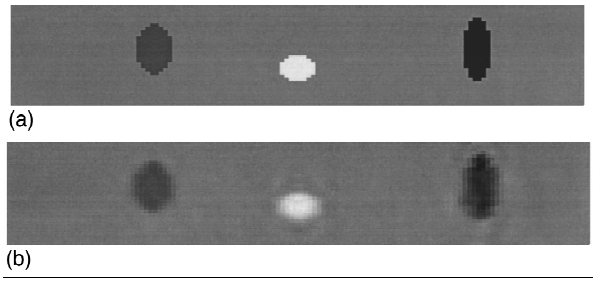


FIG. 2. (a) Gray scale of true 128 by 32 pixel phantom, $\lambda_{max}/4$ sampling was again used, where λ_{max} is the wavelength at the maximum frequency. The image space is, therefore, 32λ by 8λ in size. The center object has $\gamma = -0.2$, which corresponds to speed of sound=1878.3 m/s. The object on the right has $\gamma=0.2$, i.e., $c=1533.6$. and the object to the left (elliptical shape) has $c=1602$ m/s. There were 64 frequencies used in the simulation, which were equally spaced from 0 to $f_{max}=80$ kHz. Twelve sources were placed along the top of this image space, adjacent to it. The image space was surrounded by layers representing seawater and sedimentation, with $c_0 = 1480$ m/s, $c_1=1500$ m/s, $c_2=1680$ m/s, $c_3=2500$ m/s (see Fig. 1). (b) Reconstruction using inverse scattering of the above computer phantom.

ing low-frequency data, then to build on this using the higher-frequency information.

Parallel computers have been used to speed up the forward problems and, independently, the multifrequency calculation of the Jacobians, thereby speeding the process of functional minimization in two independent ways. The independent nature of the multifrequency calculations renders this algorithm extremely efficient when parallelized. We have used the University of Utah Supercomputer Institute's cluster of RS6000 workstations coupled with the PVMD(3) software to indicate the efficiency for the single frequency,

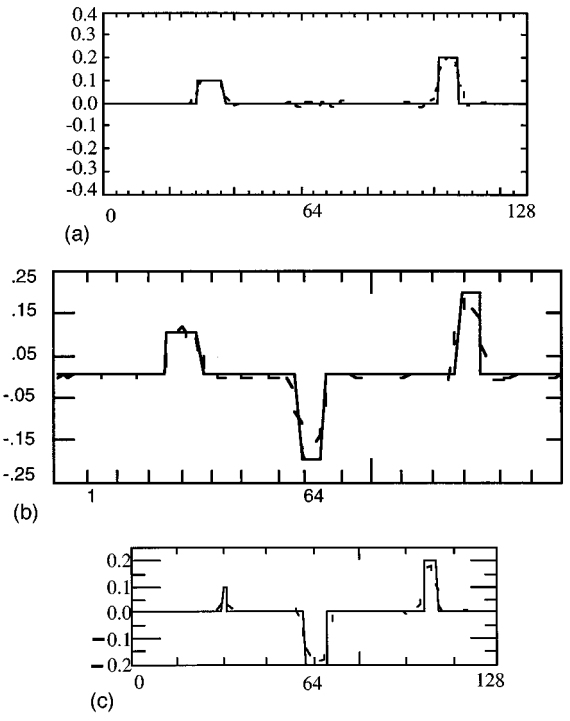


FIG. 3. Horizontal slices through image space at $z=14$ (a), and 18 (b). Dotted line is reconstructed potential, solid line is true γ. The vertical axis is γ the horizontal axis is pixel number. (c) Horizontal section through image space at $z=22$, the dotted line is the reconstruction.

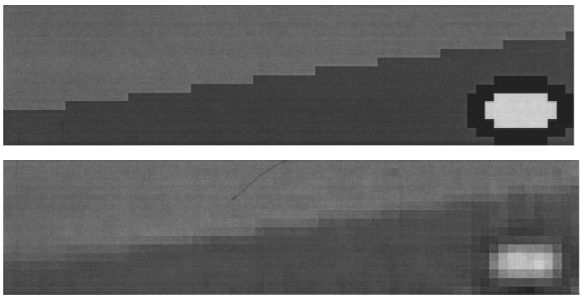


FIG. 4. There were 32 frequencies used in the reconstruction shown below the true computer phantom. These were equally spaced from $f_{max}=60$ kHz to 1/32*60 kHz. The sediment has a $\gamma=-0.195\,45$ ($c_{sediment}=1650$ m/s). The background (sea water) has a speed of sound=1480 m/s. This image space is embedded in a transition layer with $c_1=c_2=1480$ m/s (see Fig. 1). It is directly above a sediment layer with $c_3=1650$ m/s. The pixels were λ/4, where λ=1480/60=24.67 mm. The image space is 64 by 16 pixels or 16 by 4λ. The physics scales appropriately, so that $f_{max}=6$ kHz implies λ =0.247 m, and the mine is approx. 0.5 by 1.0 m, in an image space approx 3.95 m by 1 m deep. The steel shell has the corresponding horizontal section in Fig. 5.

multiple view inversion (which parallelizes in an analogous manner to the multifrequency case). These results are shown in Fig. 6.

VI. CONCLUSIONS AND FURTHER DIRECTIONS

A quantitatively accurate reconstruction of a scattering object is possible within layered media, with essentially the same algorithm as used in the homogeneous case. This ability to quantitatively reconstruct object function values (γ values) is one of the major advantages of the nonlinear reconstruction employed here over diffraction tomography, as defined in the geophysical literature. The trade-off associated with this quantitative reconstructive ability is the increased computational burden imposed. Also, as the contrast and size of the object increase, the computational burden does increase dramatically. However, the ability to reconstruct objects of higher contrast and/or size is beyond that of the Born approximation (diffraction tomography). There has been little theoretical analysis carried out to characterize the rate of convergence of the inverse problem up to this point; however, the results of this paper indicate that the reconstruction procedure is practical. The reconstruction algorithm described here can easily be generalized to the full 3-D case, and in fact this has been done and is the subject of a future

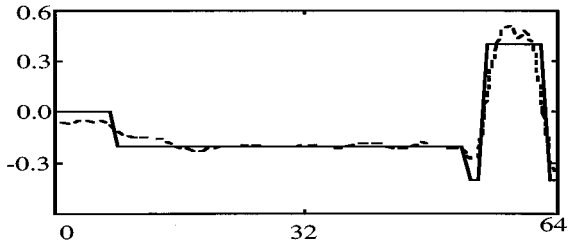


FIG. 5. The cross section of the computer phantom in Fig. 4, and its reconstruction (dotted line). The reconstruction was carried out with pure reflection mode geometry.

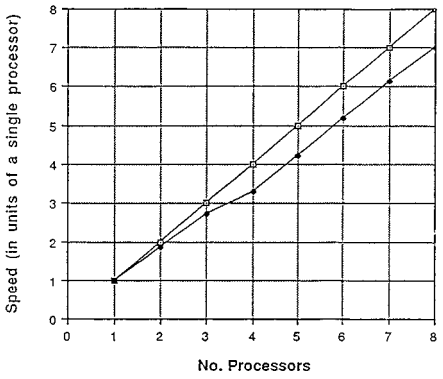


FIG. 6. Speed of cluster versus number of processors.

publication. Further generalizations that are being undertaken include inversions and forward problems in layered Biot media (poroelastic media).

It is also possible to construct similar layered Green's functions in the more general case when the layers have linear varying index of refraction, or more general smooth variation in certain special cases, for example, if the speed of sound is given by $cc'(c/c')''=\mu=$constant, where $'$ denotes differentiation. This last fact is a result of the application of the theory of Lie Group symmetries of differential equations[24] to the wave equation. The existence of an analytic solution means that the layered Green's function for this type of layered medium can be constructed, discretized, and the FFT-BCG algorithm described in this paper implemented in a manner analogous to the construction described here. The advantage of having layers which have their speed of sound with the dependence described by $cc'(c/c')''=\mu=$constant is the greater flexibility in describing a sediment with arbitrary vertical speed of sound variation with more accuracy and fewer layers than would be required if the speed of sound is required to be homogeneous within each layer.

The natural extension to three-dimensional image spaces has been carried out. The advantage of this method includes: (1) the coarse spatial frequency sampling requirements afforded by the use of the ''sinc'' basis functions and integral equations; (2) The automatic incorporation of multiple reverberations, head waves, and all refraction and diffraction effects associated with the infinite boundaries between layers above and below the image space. These head waves and other effects are incorporated into the scattered field, even though the interfaces lie outside the image grid.

There is substantial interest in the Navy for development of efficient algorithms which will accurately distinguish between buried and semi-buried (or proud) mines, and rocks, or other harmless objects on the sediment surface. The examples shown above indicate clearly that this reconstruction algorithm is capable of determining whether an object that is located by some other (say synthetic aperture) algorithm is in fact ordnance or miscellaneous/harmless material. The data collection requirements of this algorithm are admittedly severe. The requirement for accurately calibrated data (the incident field and layer parameters must be known) present not insurmountable problems to developers. The accuracy of the reconstructions indicate clearly that it is definitely worthwhile to expend energy in the direction of development of arrays capable of collecting the data required by this and related algorithms being developed at the Center for Inverse Problems in the University of Utah.

ACKNOWLEDGMENTS

We are grateful for the support provided by the U.S. Navy, NRI, Stennis Space Center, MS, in the form of a grant to the University of Utah, and to the State of Utah for a grant to the University of Utah's Center for Inverse Problems, Imaging, and Tomography, and TechniScan, Inc., Salt Lake City, UT, for a subcontract to the University of Utah under its Phase I SBIR contract from the Naval Coastal Systems Center, Panama City, FL (Grant No. N61331-94-C-0054), and the Department of the Navy, Office of Naval Research, (Ocean Acoustics Program). We have also used the University of Utah Supercomputing Institute's machine (IBM RS6000 workstation cluster) and man resources, which we gratefully acknowledge. In particular the help from Michael Pernice has been substantial.

[1] H. Gan, R. Ludwig, and P. L. Levin, ''Nonlinear diffractive inverse scattering for multiple scattering in inhomogeneous acoustic background media,'' J. Acoust. Soc. Am. **97**, 764–772 (1995).
[2] L. Misici and F. Zirilli, ''Three-dimensional inverse obstacle scattering for time harmonic acoustic waves: A numerical method,'' SIAM J. Sci. Comput. **15**, 1174–1193 (1994).
[3] D. S. Ahluwalia, G. A. Kriegsmann, and E. L. Reiss, ''Direct and inverse scattering of acoustic waves by low-speed free shear layers,'' J. Acoust. Soc. Am. **88**, 1596–1608 (1990).
[4] G. Dassios and R. J. Lucas, ''Inverse scattering for the penetrable ellipsoid and ellipsoidal boss,'' J. Acoust. Soc. Am. **99**, 1877–1882 (1996).
[5] D. N. Ghosh Roy, *Methods of Inverse Problems in Physics* (CRC, Ann Arbor, 1991).
[6] O. Lenoir, J. L. Izbicki, and P. Rembert, ''Acoustic scattering from an immersed plane multilayer: Application to the inverse problem,'' J. Acoust. Soc. Am. **91**, 601–611 (1992).
[7] D. Colton and P. Monk, ''The numerical solution of an inverse scattering problem for acoustic waves,'' IMA J. Appl. Math. **49**, 163–171 (1992).
[8] A. Zinn, ''On an optimization method for the full- and the limited-aperture problem in inverse acoustic scattering for a sound-soft obstacle,'' Inverse Probl. **5**, 239–251 (1989).
[9] S. A. Johnson, Y. Zhou, M. L. Tracy, M. J. Berggren, and F. Stenger, ''Inverse scattering solutions by a sinc basis, multiple source moment method—Part III: Fast algorithms,'' Ultrason. Imaging **6**, 103–116 (1984).
[10] D. T. Borup, S. A. Johnson, and M. J. Berggren, ''Nonperturbative diffraction tomography via Gauss–Newton iteration applied to the scattering integral equation,'' Ultrason. Imaging **14**, 69–85 (1992).
[11] C. H. Wilcox, ''Ultrasonics imaging at the AIM laboratory, University of Utah,'' Lecture presented at the IMACS International Symposium on Computational Acoustics, Harvard University, 25–28 June 1991.
[12] E. J. Ayme'-Bellegarda and T. M. Habashy, ''Forward ultrasonic scattering of multidimensional solid or fluid inclusions buried in multilayered elastic structures,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **39**, 1–10 (1992).
[13] E. J. Ayme'-Bellegarda and T. M. Habashy, ''Ultrasonic inverse scattering of multidimensional objects buried in multilayered elastic background structures,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **39**, 11–18 (1992).
[14] J. W. Wiskin, dissertation, Dept. of Mathematics, University of Utah (available from UMI (University of Microfilms International, Dissertation Services).
[15] P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (Princeton U. P., Princeton, NJ, 1968).
[16] B. Hofmann, ''Regularization of nonlinear problems and the degree of ill-posedness,'' in *Inverse Problems: Principles and Applications in Geophysics, Technology, and Medicine*, Proceedings of International Confer-

ence, Potsdam, 1993, Mathematical Research, Vol. 74, edited by G. Anger, R. Gorenflo, H. Jochmann, H. Moritz, and W. Webers (Akademie Verlag, Berlin, 1993).
[17] F. Natterer, *The Mathematics of Computerized Tomography* (Wiley, New York, 1986).
[18] F B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics*, American Institute of Physics Series in Modern Acoustics and Signal Processing (AIP, New York, 1994).
[19] M. D. Collins, ''Higher-order Pade' approximations for accurate and stable elastic parabolic equations with application to interface wave propagation,'' J. Acoust. Soc. Am. **89**, 1050–1057 (1991).
[20] J. W. Wiskin, S. A. Johnson, D. T. Borup, M. Berggren, and R. Eidens, ''Full inverse scattering versus Born-like approximation for imaging in a stratified ocean,'' *Oceans '93, Engineering in Harmony with the Ocean*, Proceedings of the IEEE Oceanic Engineering Conference, Vol. III, p. 450.
[21] G. Muller, ''The reflectivity method: A tutorial,'' J. Geophysics **58**, 153–174 (1985).
[22] F. Stenger, *Numerical Methods Based on Sinc and Analytic Functions*, Springer Series in Computational Mathematics 20 (Springer-Verlag, New York, 1993).
[23] C. C. Paige and M. A. Saunders, ''LSQR: An algorithm for sparse linear equations and sparse least squares,'' ACM Trans. Math. Soft. **8**, 43–71 (1982).
[24] G. Bluman and S. Kumei, ''Exact solutions for wave equations of two-layered media with smooth transition,'' J. Math. Phys. (N.Y.) **29**, 86–96 (1988).

Matched-field evaluation of acoustic scattering from Arctic ice

Tarun K. Kapoor and Henrik Schmidt
Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139

(Received 22 September 1995; accepted for publication 1 April 1997)

This paper investigates the three-dimensional acoustic scatter from large-scale features under the Arctic ice cover using field data collected during the CEAREX experiments in April 1989. The analysis involves the identification and isolation of the features under the ice canopy, and the simultaneous evaluation of their spatial scattering characteristics. A two-step matched-field processing (MFP) algorithm is used to solve this complex multiparameter estimation problem. Using one large ($\sim$1200 m) and four short ($\sim$300 m) vertical line arrays, and a crossed horizontal hydrophone array in a matched-field processing scheme, beamforming under the ice sheet was performed to obtain an approximate map of the rough interface between ice and water. This analysis yields estimates for scattering strengths for frequencies ranging between 25 and 105 Hz. It is shown that scatter from the elastic ice surface is a frequency selective phenomenon. This beamformed map is then used to isolate a single discrete scatterer, and subsequently, its scattering characteristics are evaluated. Simultaneously, an idealized environmental model is developed to compute scatter from a canonical protuberance. Matching the theoretical results from the analytical model with the experimental data, which consists of contributions from multiple scatterers, yields the statistics of the distribution of ice protuberances. © *1997 Acoustical Society of America.*
[S0001-4966(97)00608-5]

PACS numbers: 43.30.Gv, 43.30.Ma, 43.30.Hw [JHM]

INTRODUCTION

Scattering of sound waves from the underside of ice is a problem of interest to researchers in Arctic acoustics. The Arctic Ocean waveguide, with its upward refracting sound velocity profile (SVP), causes any long-range propagating signal to repeatedly interact with the ice cover, thereby changing its characteristics. An understanding of this interaction between acoustic waves and the rough elastic interface is essential in improving the modeling of long-range propagation in the Arctic Ocean. The Arctic ice canopy contains a wide spectrum of roughness, extending from the minute deviations from the plane surface to features whose dimensions are much larger than the wavelengths of the acoustic signals typically used for long-range propagation in the Arctic Ocean ($f<100$ Hz). Recent investigations[1–3] have considered scattering from small scales of roughness ($ka\ll 1$) by using the method of small perturbations, while for large-scale scatterers ($ka\gg 1$), Kirchoff tangent plane approaches are appropriate.[4] In the intermediate regime [$ka\sim O(1)$], where the dimensions of the scatterers are comparable with the wavelength of the acoustic waves, no traditional modeling tools apply.

One of the earliest attempts by researchers in this field was to model these features as, for example, a hemispherical protuberance on an infinite plane.[5,6] The main drawback of this *boss* model approach is that it ignores the elasticity of the rough surface. Moreover, it provides no insight into the physical processes and wave mechanisms involved in the scattering of sound from ice. Recently, some researchers have modeled the scattering from a two-dimensional fluid–elastic interface using the finite difference (FD)[7] and boundary element (BEM)[8] approaches. Regarding three-dimensional scattering, there have been attempts using FD[9] and a boundary integral formulation.[10] However, due to computational limitations, their analyses were restricted to scattering from a compact feature on a rigid boundary.

In this paper, we adopt an alternate approach of analyzing field experimental data, and simultaneously develop a theoretical model to aid the evaluation of three-dimensional scatter. Even though the idealized environmental model presented here does not represent the real scattering scenario, it is more complete than such earlier representations in terms of the fundamental scattering mechanisms involved. We consider the scattering of low-frequency (25–105 Hz) acoustic waves from the Arctic ice sheet via matched-field processing of short-range reverberation returns. The experimental data that we analyzed were collected on 15 April 1989 during the Coordinated Eastern Arctic Experiments (CEAREX) conducted in the Norwegian–Iceland Sea. The receiver array geometry consisted of a crossed horizontal array deployed at a depth of 60.0 m, one 1200-long-m-long vertical line array, and four short vertical arrays, each of length 300 m. The sources used in the experiments were 1.8-lb SUS charges detonated at nominal depths of 244 m. Recently,[11] some results have been published from the analysis of reverberation data collected on the long vertical line array using plane-wave beamforming. The results presented here are based on matched-field processing using a point source replica field by combining the outputs from all the arrays.[12]

We begin with a statement of the problem in Sec. I, followed by a brief description of the CEAREX experiments in Sec. II. In Sec. III, we carry out matched-field processing of the reverberation data to image the undersurface of the ice which allows the identification and isolation of

strong scatterers. Next we perform an analysis of the scattering characteristics from a single discrete scatterer in Sec. IV, and, matching the experimental data with the analytical model, we compute the statistics of the distribution of protuberances under the ice canopy.

I. THE PROBLEM

The focus of this paper is to analyze scattering from a single three-dimensional feature, or protuberance under the Arctic ice canopy using field experimental data. In its most general form, this is a complex multiparameter estimation problem with the unknown quantities being: (i) the location of the protuberances; and (ii) their spatial scattering characteristics. Among the multitude of methods available to solve this problem is the matched-field processor (MFP), which maximizes the correlation between data and model.[13] Within the gamut of matched-field estimators are included the conventional and high-resolution processors. While the conventional processor uniformly weights the dependence on the search parameters, the high-resolution processor adaptively adjusts the weights so as to maximize performance. High-resolution estimates of these unknown parameters may be obtained using this processor, for example, the maximum likelihood method (MLM). However, it is severely constrained by requiring an exact knowledge of the environment that defines the model or replica fields. On the other hand, the performance of the conventional processor (e.g., the Bartlett processor) is relatively insensitive to environmental mismatch. Since we will be analyzing real field data, with its associated uncertainties in positions of source and receiver arrays, the performance of the MLM processor will rapidly deteriorate, producing erroneous results. Therefore, we shall instead carry out our analysis with the Bartlett processor.

In general, the scattered field from a large-scale feature under ice may be represented by $g(\rho,f,\mathbf{r}_p,\mathbf{r}_0)$, where ρ is the size of the feature, f is the frequency, and $\mathbf{r}_p$ and $\mathbf{r}_0$ are the three-dimensional coordinates of the location of the protuberance and the receiver. Note that we have no *a priori* information regarding the functional dependence of g on these parameters. This is inherently a complex problem to solve where we need to simultaneously determine the location of the feature and its scattering characteristics. Therefore, we make this analysis tractable by solving this problem in two stages, where the first step involves isolation and identification of discrete scatterers under the ice cover. Our theoretical analysis in Refs. 14 and 15 guides us in assuming that the three-dimensional features resemble point radiators with a quadrupolar scattering pattern. Moreover, Fricke[7] has demonstrated that the scattering pattern from a two-dimensional protuberance on an elastic plate, i.e., an elastic keel, resembles a deformed quadrupole. This assumed scattering pattern will then aid in providing an approximate map of the undersurface of the rough ice cover. We can then proceed to identify and isolate discrete features under the ice canopy. Subsequently, for any isolated protuberance, we can evaluate its spatial scattering characteristics using the reverberation data received on the horizontal and vertical arrays.

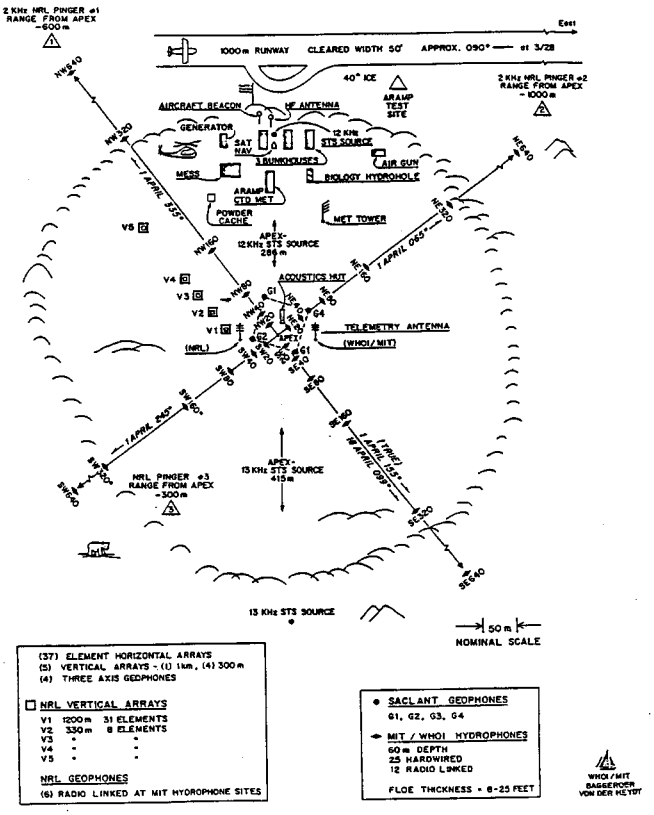


FIG. 1. CEAREX acoustics camp layout. North points upward and East points to the right. (Adapted from WHOI technical memorandum; see Ref. 16)

II. THE CEAREX REVERBERATION EXPERIMENTS

Figure 1 shows a schematic of the layout of the acoustics camp (A-camp) during the CEAREX experiments. The camp was situated on a circular ridge about 650 m in diameter, as shown by the circular boundary in Fig. 1. Geographical north points upward and east points to the right. The acoustics hut is located near the center of the camp floe which is also the apex of the horizontal hydrophone array. The main camp is located about 300 m due north. Also seen in the pictorial representation, about 250 m north and 30 m east of the apex of the array, is Wiebe's biology hydrohole through which SUS charges for the reverberation experiments were dropped. Aerial pictures of the camp revealed that the ice around the camp had a rough topography.

Two sets of arrays were deployed during the reverberation experiments. The first consisted of a crossed horizontal hydrophone array with unequally spaced sensors (Mill's Cross), located at nominal depths of 60 m, and was monitored jointly by the Massachusetts Institute of Technology (MIT) and the Woods Hole Oceanographic Institution (WHOI). The second receiver array system was deployed by the Naval Research Laboratory (NRL), and consisted of one long ($\sim$1200 m) and four short ($\sim$300 m) vertical arrays. These two sets of arrays were connected to different data acquisition systems. The MIT/WHOI data recording system acquired the digital data at a sampling rate of 1 kHz, while the NRL data acquisition system recorded the data at a sampling frequency of 689 Hz.

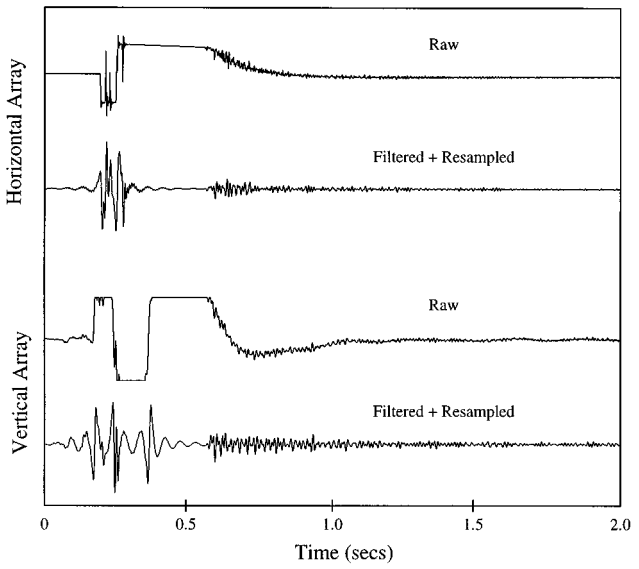


FIG. 2. Time series data received at the horizontal and vertical arrays. The figure shows examples of typical raw time series recorded along with their filtered and resampled versions.

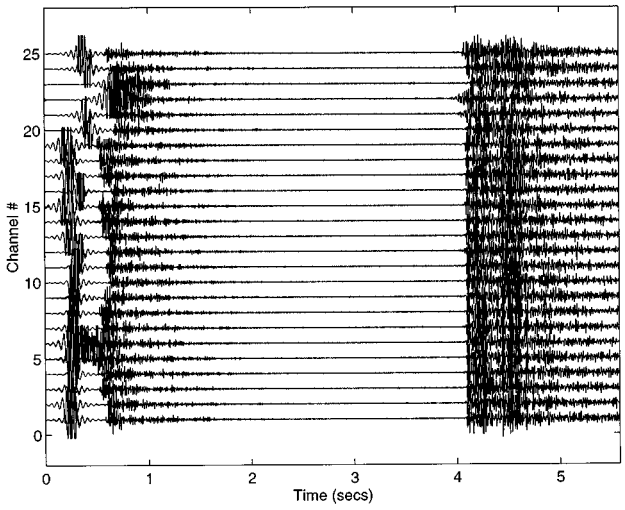


FIG. 3. Filtered and resampled data from the Mill's Cross horizontal line array.

The crossed horizontal hydrophone array consisted of 25 logarithmically spaced hydrophones distributed along the two legs of the array, as depicted schematically by the black diamonds in Fig. 1. The hydrophones were placed at nominal distances of 20 m, 40 m, and so on up to 640 m from the apex of the array. The northeast–southwest (NE-SW) and the northwest–southeast (NW-SE) legs were thus approximately 1280 m in length. The positions of these hydrophones were monitored in real time via an acoustic sensor tracking system (STS). Six high-frequency pingers were deployed at calibrated positions, and periodic travel time measurements from these to the various hydrophones were used in a simulated annealing algorithm to estimate sensor positions.[16]

The 1200-m-long vertical line array consisted of 31 hydrophones, 21 of which were spaced 30 m apart at nominal depths of 30–630 m, and 10 of which were spaced 60 m apart at nominal depths of 690–1230 m. The four short 300-m-long vertical line arrays consisted of 8 hydrophones each at nominal depths of 30, 90, 120, 150, 180, 210, 270, and 330 m. The fixation points of the top of the five vertical arrays were approximately along the NW leg of the horizontal array as shown in Fig. 1. Six hydrophones on the long vertical array and five hydrophones on the short arrays had severe 60-Hz contamination[17] and thus were not used in the analysis. Using a high-frequency pinger, positions for only 16 channels on the long vertical array were recorded. Only the locations of the fixation points were recorded for the shorter arrays.[17] The positions of the missing hydrophones on the long array were interpolated from its displaced shape. The sensor positions of the shorter arrays were deduced by assuming a displaced shape similar to that of the longer array, and by noting the arrival times of the direct pulse. The details of this analysis have been presented in Ref. 15.

Five shots were fired at 15-min intervals, of which one turned out to be a dud. Thus data from only four shots were available for subsequent processing. Figure 2 shows examples of the typical raw time series recorded for the 25 channels of the horizontal array and the 31 channels of the long vertical array, respectively. Each hydrophone of the horizontal array saturates with the arrival of the direct shock pressure pulse. After a characteristic system decay time, the sensors come back on line, as shown in Fig. 2. On the other hand, the arrival of the shock pressure pulse at the hydrophone nearest to the source results in a simultaneous saturation of all the hydrophones on the vertical arrays.

Before any data could be processed, the source detonation location had to be estimated, followed by synchronization of the two different data sets to the same time frame of reference. The bubble pulses are easily observable in the raw time series for the horizontal array, and this information was used to determine the actual depth of detonation of the SUS charges. The source locations were estimated by noting the arrival time of the direct pressure pulse at the hydrophones, and using these in a least-squares estimation procedure.[15] Only the horizontal array was used since it provides a much better resolution in estimating the two-dimensional range. Next, the MIT/WHOI and NRL data were filtered and resampled to the common sampling rate of 520 Hz. The resulting time series were then bandpass filtered between 19.5 and 130 Hz and, subsequently, the gains and sensitivities were also corrected for the two different data sets.

Figure 3 shows the conditioned time series at the horizontal array hydrophones after filtering and resampling. The direct shock pressure pulse is seen to arrive around 0.2 s at most of the receivers on the horizontal array. The ocean bottom is roughly at a depth of 3 km, and therefore, as we would expect, the first bottom arrivals come in about 4 s later, i.e., at a time of 4.2 s. Also seen are the surface-reflected bottom returns that arrive about 0.2 s later. The usable data for estimating short-range scattering from under the ice surface are then contained between 0.7 s, i.e., after the sensors come back on line, and 4 s. On the other hand, as shown in Fig. 4, the direct arrival is seen to come in at around 0.2 s for the topmost receiver of the long vertical array, denoted by Channel # 1. The receiver time series are

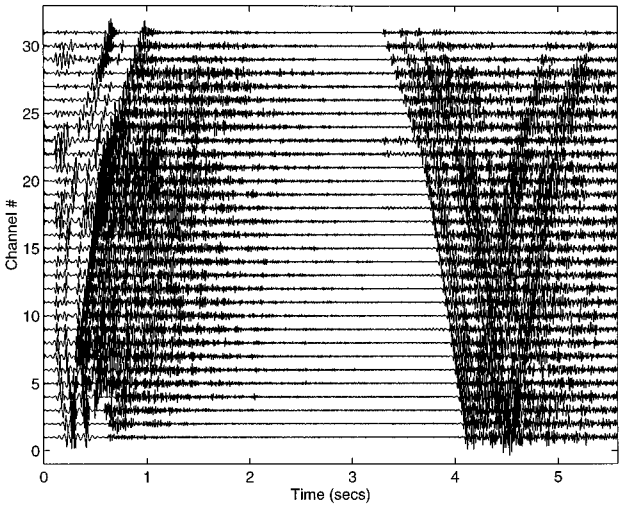


FIG. 4. Filtered and resampled data from the 1200-m-long vertical line array.

stacked in the order of depth with the deepest hydrophone corresponding to Channel # 31. Note how the direct bottom returns and the surface-reflected bottom returns are clearly separated for the deeper receivers of the long vertical array.

III. MATCHED-FIELD REVERBERATION ANALYSIS

The matched-field processor (MFP) correlates the field arriving at the array with model replicas of the expected field to find the replica field that best matches the received signal. Matched-field processing techniques have traditionally been used by researchers in acoustics in the past in the context of plane-wave beamforming. However, recent literature[13] has demonstrated the applicability of this approach in the context of localizing a distant source in range, depth, and bearing. Here, we adapt this algorithm to near-field beamforming or focusing.

A. Near-field beamforming

For the purpose of illustration, consider a time-harmonic point radiator of strength P_0 and frequency ω located at (x_s, y_s, z_s) in the near field of an array of receivers of arbitrary geometry, as shown in Fig. 5. Assuming the radiation pattern to be denoted by $B(\omega, \theta, \varphi)$, the field received by the N sensors of the array, located at (x_j, y_j, z_j), may be expressed as

$$P_j(\omega) = P_0(\omega) B(\omega, \theta_j, \varphi_j) \frac{e^{ikR_j^T}}{R_j^T} + \mathcal{N}_j(\omega), \quad j = 1, \ldots, N. \tag{1}$$

Here $P_0(\omega)$ is the source strength at 1 m, and θ_j and φ_j are the polar and azimuthal angles, respectively, subtended by the source at the jth receiver. Here, $\mathcal{N}_j(\omega)$ is the noise term uncorrelated from sensor to sensor, k is the acoustic wave number with $k = \omega/c$, and c is the speed of sound in the medium. The distance from the source to the jth receiver is given by

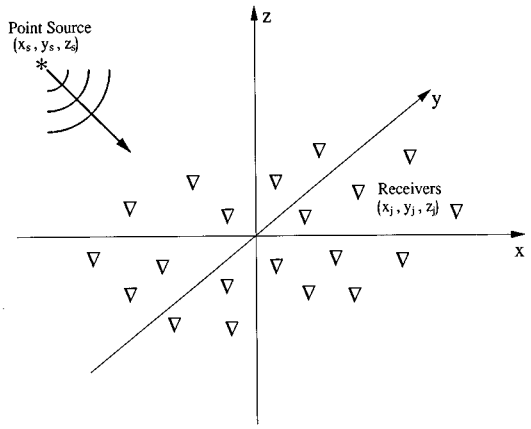


FIG. 5. Receiver array and source geometry for near-field beamforming or focusing.

$$R_j^T = [(x_s - x_j)^2 + (y_s - y_j)^2 + (z_s - z_j)^2]^{1/2}, \quad j = 1, \ldots, N. \tag{2}$$

The MFP algorithm or near-field beamforming operation estimates the strength of the point source and its location from the received signal at the N sensors. Since the source is modeled as a point radiator, the signal strength may be estimated by exploiting the difference in-phase of the spherical wavefront at each of the sensors. For a trial source position (x, y, z), the output of the near-field beamformer is given by

$$\hat{P}(\omega) = \sum_{j=1}^{N} W_j R_j \mathcal{B}_j P_j(\omega) e^{-ikR_j}, \tag{3}$$

where

$$R_j = [(x - x_j)^2 + (y - y_j)^2 + (z - z_j)^2]^{1/2}, \quad j = 1, \ldots, N, \tag{4}$$

and

$$\mathcal{B}_j = \frac{b_j}{\Sigma_{j=1}^{N} b_j}, \quad b_j = \frac{1}{B(\omega, \theta_j, \varphi_j)}. \tag{5}$$

Here, W_j is the weighting or array taper vector, and the radial distance R_j appears in the numerator to account for the spherical spreading loss. The term $B(\omega, \theta_j, \varphi_j)$ accounts for the source radiation pattern and has been normalized to obtain a 0-dB response at the maximum response axis (MRA). Then, Eq. (3) will yield the estimate of the source strength when the trial source position coincides with the true source coordinates.

B. Adaptive focusing for CEAREX arrays

The analysis presented here combines the outputs from the horizontal array, the long vertical line array, and the four short vertical arrays. The schematic geometry of the source, receiver arrays, and beamforming patches is shown in Fig. 6. The response of the simulated volumetric array was computed by positioning a broadband point source at the location (900,900,3.4), i.e., in the center of the NE patch in Fig. 6. Moreover, the value $z = 3.4$ m corresponds to the mean depth of the lower surface of the ice sheet assumed to be uniform.

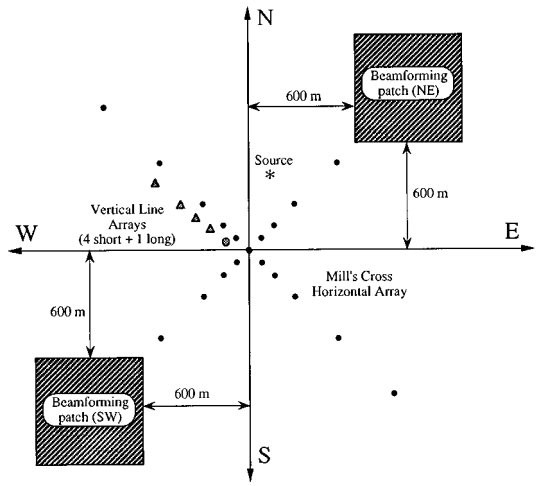


FIG. 6. Relative geometry of source, receiver arrays, and beamforming patches.

Next, a synthetic time series was generated for all the sensors by propagating the field using spherical spreading and adding white Gaussian noise, with an effective signal-to-noise ratio (SNR) of 60 dB. The output from the three arrays is combined adaptively to optimize the performance of the volumetric array in both range and azimuth. After an extensive study of various possible combinations of the conventional beamformer outputs, it was found that the best response is obtained by combining the outputs from the three arrays—horizontal, long vertical, and short vertical—in the ratio 0.6, 0.2, 0.2, respectively. An unbiased estimate is obtained by requiring that the sum of the three factors equal 1. Figure 7 shows the ambiguity function for the volumetric array obtained by focusing, where we have also geometrically averaged the output of the processor over the frequency band 45–55 Hz. It has been shown previously[13] that frequency averaging the beamformer output further reduces the sidelobe levels. With increasing frequency, similar beampatterns are obtained, but with higher resolution in range.

The geometric paths from source to focusing points under the ice, and then on to the receivers, were determined using ray tracing in a bilinear approximation of the sound velocity profile (SVP). The scattering strength calculations are based on the standard sonar equation

$$\mathrm{SS} = L_{RL} + H_{S\to FP} + H_{FP\to R} - L_{SL} - 10\log_{10} A, \qquad (6)$$

where L_{RL} is the received pressure level at the receivers in units of dB, $H_{S\to FP}$ is the transmission loss from source to focusing point, and $H_{FP\to R}$ is the transmission loss from the focusing point to the receiver in dB. L_{SL} is the source spectral level in dB at a reference distance of 1 m, and A is the ensonified area. The source spectral levels were deduced from the recent work by Chapman.[18] Equation (6) yields an estimate of the scattering strength SS in units of dB/m^2.

The algorithm for computing the scattering strength from the received reverberation time series is as follows. For each given sensor, we compute the two-way travel time from source to focusing point, and then on to the receiver using ray tracing. For a constant SVP, the loci of points on the focusing plane for a fixed two-way travel time correspond to an ellipse, as shown in Fig. 8. This is also true for the varying SVP, provided we determine the travel times from correct ray paths. We centered a Hamming window, 101

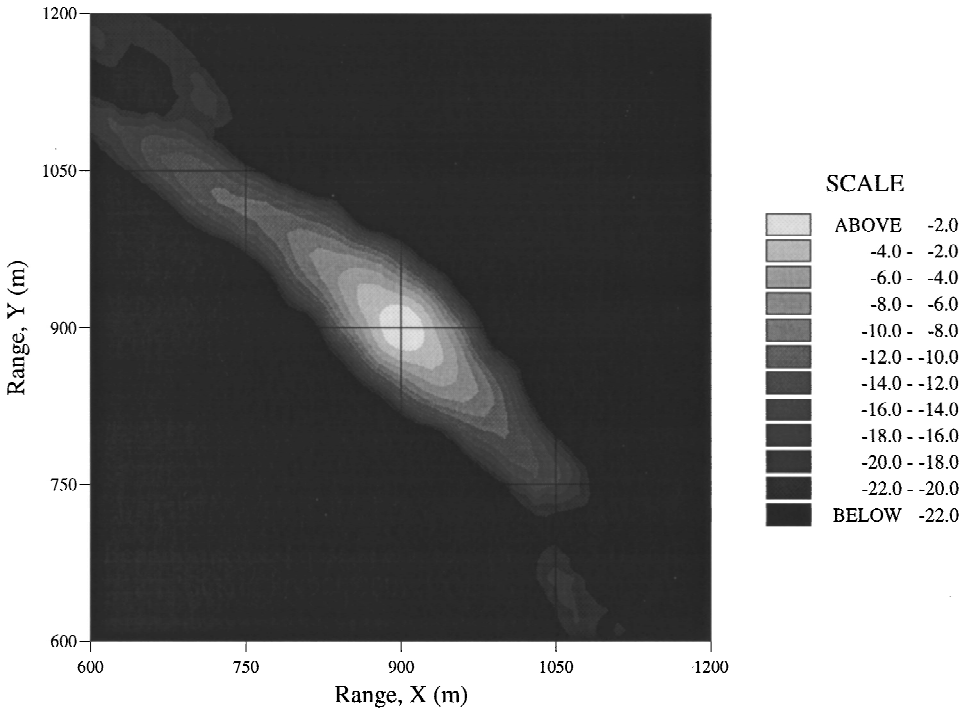


FIG. 7. Beamformer response of the simulated volumetric array averaged over the frequency band 45–55 Hz for a source positioned at (900,900). Note the circular ambiguity in beampattern.

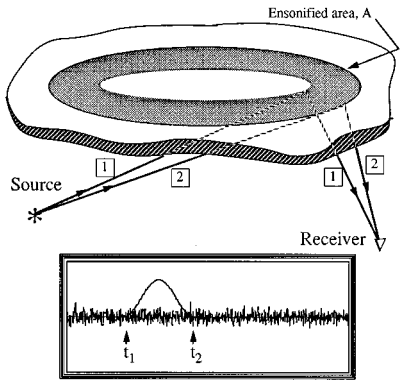


FIG. 8. Ensonified area for the source, receiver, and focusing point configuration. Here t_1 and t_2 are the two-way travel times for paths 1 and 2.

samples or 0.194 s long, around the two-way travel time corresponding to the focusing point of interest on each hydrophone. The area of intersection between the two ellipses corresponding to the end points of the windowed reverberation time segment is then the ensonified area. Coherent summation across channels of the suitably normalized, windowed outputs in the time domain for each array separately yields the beamformer output as

$$\hat{P}^i = \sum_{j=1}^{N^i} \frac{1}{\sqrt{A_j^i}} W_j^i \mathscr{B}_j^i R_{j,s}^i R_{j,r}^i P_j^i(t_{1,j}:t_{2,j}). \tag{7}$$

Here, $\hat{P}^i$ is the coherent beamformer output, N^i is the number of receivers in the ith array, and W_j^i is the window for the jth sensor in the ith array. Also, $R_{j,s}^i$ is the ray path length from source to focusing point, and $R_{j,r}^i$ is the ray path length from the focusing point to the receiver. $P_j^i(t_{1,j}:t_{2,j})$ is the segment of data 101 samples long, centered at t_c for the jth sensor in the ith array, and A_j^i is the ensonified area for each sensor.

The sensor outputs have also been weighted by the suitably normalized assumed radiation pattern $\mathscr{B}_j^i$. Our theoretical analysis[14,15] establishes that the scatter from an isolated three-dimensional protuberance on a fluid-loaded plate is symmetric with respect to the azimuthal coordinate φ. It has also been shown[7,15] that the polar radiation pattern from a discrete feature under ice resembles that of a deformed quadrupole in the low-frequency regime. Motivated by these analyses, we approximate the radiation pattern from these ''hot spots'' as

$$B(\omega,\theta,\varphi) = \cos\,\varphi\,\sin 2\theta, \tag{8}$$

where φ and θ are the azimuthal and polar radiation angles, respectively. In other words, the replica field for the matched-field processing scheme is that due to a quadrupolar radiator. This radiation pattern must then be normalized to obtain an unbiased estimate. Following Eq. (5), this is done by defining

$$\mathscr{B}_j^i = \frac{b_j^i}{\Sigma_{j=1}^N b_j^i}, \quad b_j^i = \frac{1}{\cos\,\varphi_j^i\,\sin 2\theta_j^i}. \tag{9}$$

As the beampattern is assumed to be constant over this frequency regime, we can therefore take advantage of time-domain beamforming.

The beamformer output of the three arrays is then combined adaptively as

$$P = \sum_{i=1}^{3} v_i \hat{P}^i, \tag{10}$$

where v_i is relative weighting of the outputs from the three arrays with $\Sigma_{i=1}^3 v_i = 1$. Finally, the scattering strength is found by transforming the beamformer output to the frequency domain, averaging the reverberation level in dB over 10-Hz frequency bands, and subtracting the source level at 1 m, i.e.,

$$\mathrm{SS}(f_c) = \left[\frac{1}{f_2 - f_1}\sum_{f=f_1}^{f_2} L_{RL}(f\,)\right] - L_{SL}(f_c). \tag{11}$$

The shaded rectangular patch in Fig. 6, NE of the apex of the array, is the area on which we focused to estimate the scattering strength. Ideally, it would be preferable to look broadside to the horizontal array, i.e., vertically up. However, this was not possible due to the limitation of the available array geometry. As pointed out before, the arrivals from features broadside to the array come in at times close to the direct path arrivals, and hence are clipped. Therefore, the analysis was confined to features at least 850 m from the apex of the array.

Using this adaptive focusing methodology, we focus on an area located approximately 850 m NE of the apex of the array, as shown in Fig. 6. The patch of ice is 600 m square, and we have computed the scattering strength at uniform intervals of 10 m. Results for the frequency band 45–55 Hz are shown in Fig. 9, where we have averaged the mean-square values of scattering strength from the four different shots. The horizontal and vertical axes correspond to the range from the apex of the horizontal array. We can clearly observe the ambiguity pattern and sidelobes of our array processing scheme superposed in the plot. We observe strong scatterers at approximate grid points of (980,890), (1000,1040), (750,1130), and (720,680). Our resolution of these discrete features is limited by the performance of the volumetric array. Therefore, it is not possible to directly infer the length scales of the scatterers from this plot due to the wide beamwidth of the mainlobe. However, strong scatter in this frequency range would suggest length scales of $L \sim \lambda$, or $L \sim 30$ m. Moreover, there could possibly be many small features within the mainlobe of this ambiguity pattern, and those whose sizes match the wavelengths of the acoustic field will be the ones that scatter with the highest magnitude. However, note the strong scatterer at (980,890). We will return to this isolated feature later in our analysis, when we compute its scattering response as a function of grazing angle of scatter.

Figure 10 shows the scattering strength for the same ice patch for a lower-frequency band of 25–35 Hz. Once again, we have presented the mean results from four different shots. Note that this plot is qualitatively similar to Fig. 9. However, at this frequency, the resolution in range is approximately 50

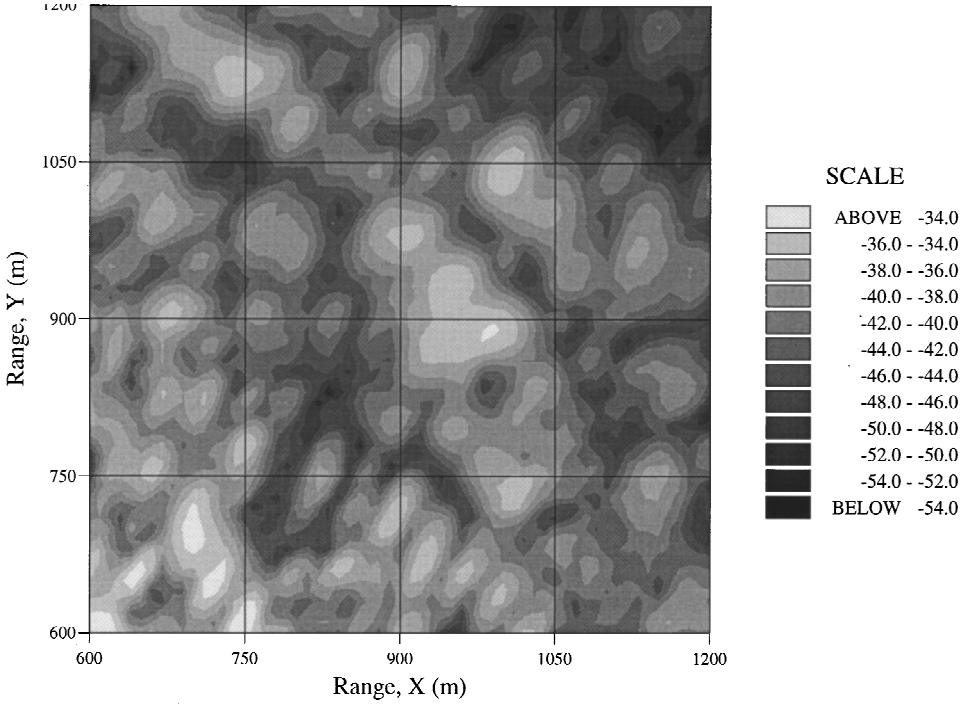


FIG. 9. Contours of constant levels of mean scattering strength for the patch of ice located NE relative to the apex of the array, and averaged over the frequency band 45–55 Hz.

m. Correspondingly, there is increased uncertainty about the size of the scatterers. Note that the levels observed here are about 4–6 dB lower than those in the 45–55 Hz band.

The striking difference between the 30- and 50-Hz results is that features at a given location do not scatter as strongly in different frequency bands, i.e., features that scatter strongly in the 50-Hz band do not scatter as strongly in the 30-Hz frequency band. In other words, the bright spots

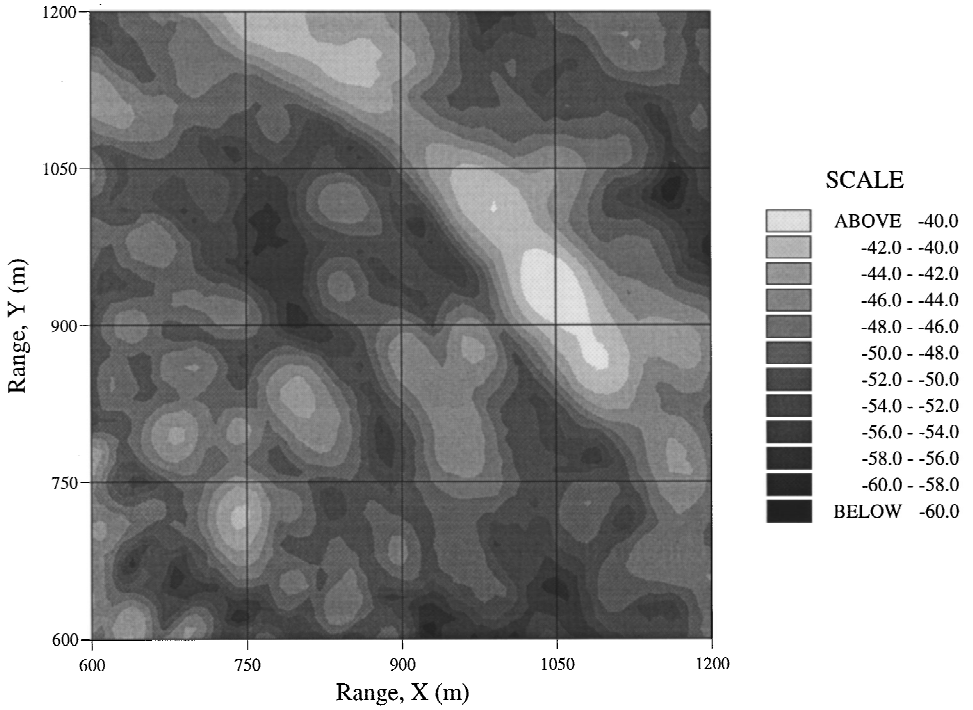


FIG. 10. Contours of constant levels of mean scattering strength for the patch of ice located NE relative to the apex of the array, and averaged over the frequency band 25–35 Hz.

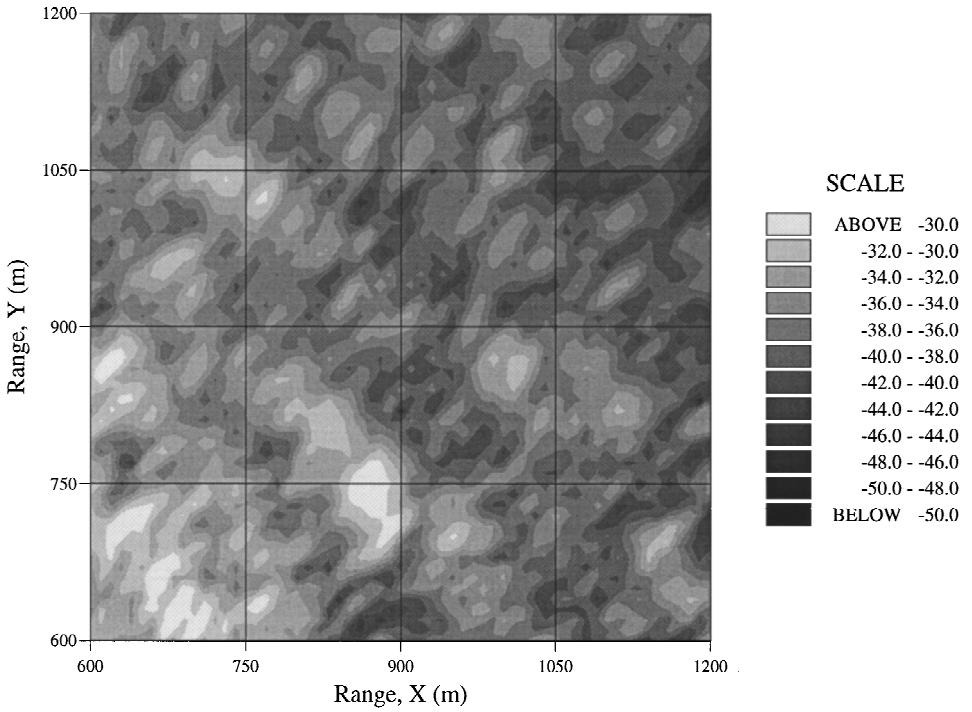


FIG. 11. Contours of constant levels of mean scattering strength for the patch of ice located NE relative to the apex of the array, and averaged over the frequency band 75–85 Hz.

do not show up at exactly the same place in the two plots. This can also be explained by recalling that within the main lobe of the ambiguity function there lie many features of varying sizes and orientation. At a given frequency of the incident acoustic waves, the features whose scattering characteristics match those of the incident field radiate most strongly. Therefore, we infer that scatter from ice is a frequency selective process.

Figure 11 shows the scattering strength for the same patch of ice for a higher-frequency band of 75–85 Hz. Once again we observe qualitative similarities with the results presented earlier for the lower-frequency bands. Our previous deduction about frequency selectivity is reiterated by these results. For example, in the 80-Hz results, we observe a strong scatterer at (880,740) which is not as prominent in the 30- and 50-Hz results. However, we continue to observe increasing levels of scatter with increasing frequency. Therefore, as postulated before, these results suggest a strong frequency dependence.

It is also informative to look at the scattering patterns from other areas under the ice canopy. This is essential to demonstrate that the conclusions at which we have arrived for the patch NE of the apex of the horizontal array are also valid for other areas under the ice cover. Therefore, we consider another region of ice 850 m SW of the apex of the array. This patch is diagonally opposite to the NE patch considered previously. The scattering patterns for two frequency bands centered at 30 and 50 Hz are presented in Fig. 12 and Fig. 13, respectively. Once again we observe the circular ambiguity pattern of the volumetric array. Also, the qualitative differences about strong features that we mentioned for the NE patch are also observed here. Moreover, consistent with the results for the NE patch, the scattering levels also increase with frequency.

These results are a little deceptive in the sense that they show that features show up in the same place in both frequency bands. For example, consider the bright feature in the 50-Hz band at $(-650,-850)$. It seems that it appears at the same location in the 30-Hz results. However, in the 30-Hz frequency band we observe two peaks surrounding this location. Neither of these necessarily correspond to the feature observed in Fig. 13. As another example consider the feature at $(-825,-900)$ in the 30-Hz band. There is no corresponding bright spot in the 50-Hz results. One can easily pick out other differences in the locations of these ''hot spots'' in the two plots. Therefore, we can conclusively state that scatter is indeed a frequency selective process.

IV. SCATTERING FROM A DISCRETE PROTUBERANCE

Having obtained a rough map of the undersurface of the ice cover, we can then proceed to the next step of the global estimation problem and determine the spatial scattering characteristics of individual discrete features. For the purpose of illustration, we present analysis for only one single feature which can be identified and isolated from our results in the previous section. This stage includes the development of an idealized environmental model for comparison with results from the analysis of experimental data. Finally, matching the

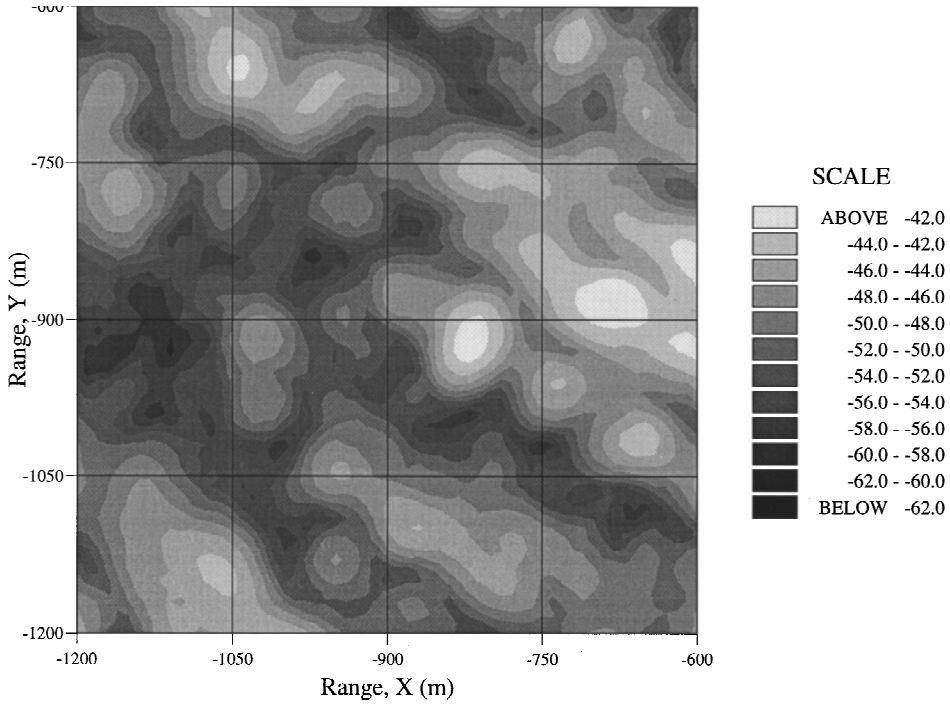


FIG. 12. Contours of constant levels of mean scattering strength for the scattering patch located SW of the apex of the horizontal array, and averaged over the frequency band 25–35 Hz.

theoretical results for a single feature with the experimental data which consists of contributions from many scatterers, we invert for the statistics of the distribution of ice protuberances.

A. Analytical model of protuberance

We begin by constructing an analytical model of scattering from a typical large protuberance under the ice sheet.

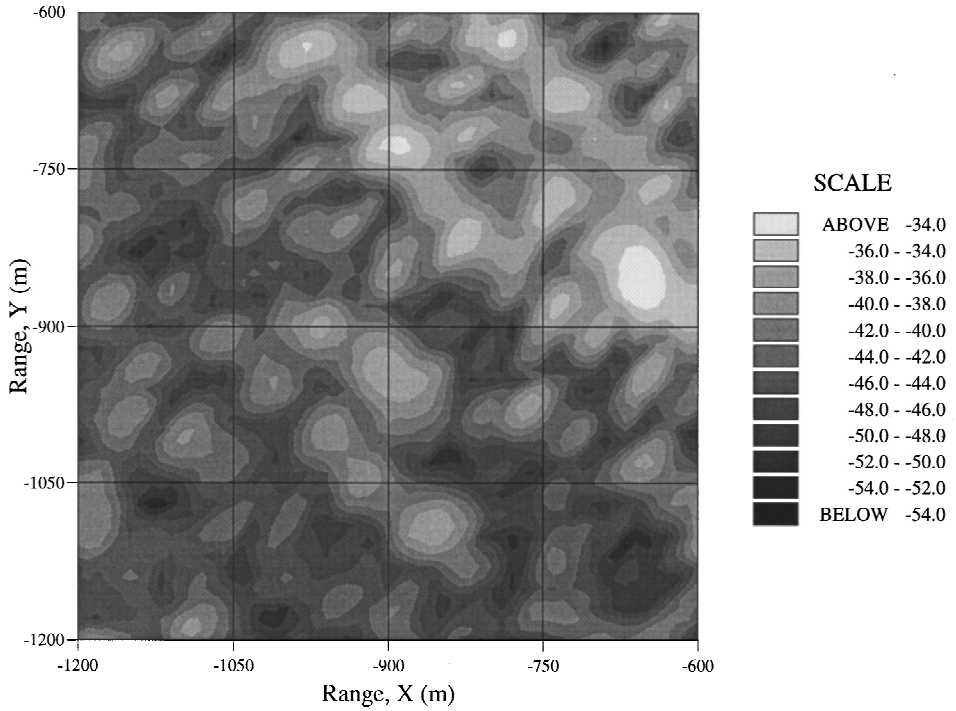


FIG. 13. Contours of constant levels of mean scattering strength for the scattering patch located SW of the apex of the horizontal array, and averaged over the frequency band 45–55 Hz.

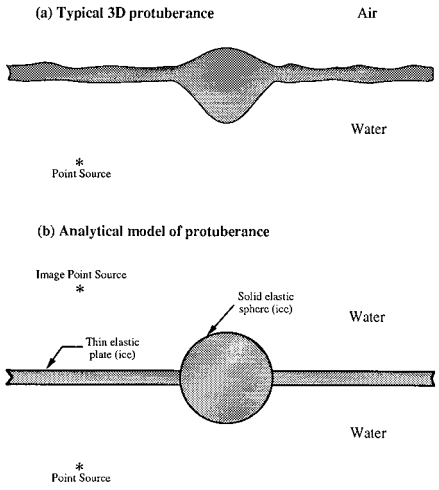


FIG. 14. (a) Pictorial representation of a typical large three-dimensional feature under the ice sheet. (b) Analytical model of the protuberance. The pressure release surface at the ice–air interface is approximated by including an image point source.

Previous analyses which have modeled these as, for example, a hemispherical protuberance on an infinite plane[5,6] failed to account for the elasticity of the ice plate. As a significant improvement in modeling the three-dimensional feature, our prototype is a solid elastic sphere attached to a thin elastic plate along its equatorial plane. Figure 14(a) shows the pictorial representation of a typical large-scale feature, or ''hot spot,'' under the Arctic ice canopy. Note that the ice plate is not of uniform thickness, and may be comprised of a multitude of such features, i.e., many smaller features cemented together giving the appearance of one large feature. In Fig. 14(b) we have shown the idealized model used in our analysis to represent this protuberance. The details of the methodology for modeling scattering from this completely submerged coupled model have been previously presented by Kapoor and Schmidt.[14] However, we need to modify the model presented there to account for the free surface at the ice–air interface. Assuming that the air–ice interface can be approximated as a free surface by including a negative image source, as shown in Fig. 14(b), we write the total scattered field as

$$p_t = p_r - p_i . \tag{12}$$

Here p_t represents the total scattered field, p_r is the scatter from the plate–sphere coupled structure due to the real source located at $(r_0, \theta_0, \varphi_0)$, and p_i is the scattered pressure due to the image source located at $(r_0, \pi - \theta_0, \varphi_0)$. Evidently this is not an exact solution because of the excitation of elastic waves in the plate and the sphere. Inclusion of a negative image source does not cancel the stresses at the upper hemispherical surface.[15] A more exact analysis requires the application of a stress distribution on the entire upper surface such that all resulting stresses vanish there. This involves an integral equation formulation and, for the purpose of our analysis here, will not be carried out. However, this idealized environmental model may be viewed as a first-order approximation in representing the real scattering scenario.

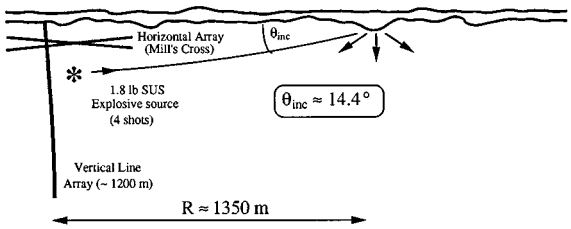


FIG. 15. Experimental geometry of source, protuberance, and long vertical and crossed horizontal arrays.

B. Experimental data analysis

We now return to the analysis of the experimental data to determine the scattering characteristics of a single, isolated feature under the ice canopy. As mentioned earlier, for the purpose of illustration, the focal point of this part of the analysis will be the discrete scatterer isolated earlier at the grid point (980,890). The sensors of horizontal array are located at a mean depth of 60 m, and therefore, the polar angle of scatter at all the sensors is clustered around a very small angle, with a mean value of 5°. Thus reverberation data from the long vertical array only are used to determine its polar (vertical) radiation pattern.

The point source is located at a range of 1150 m, and is incident on the scatterer at a grazing angle of 14.4°. The vertical array has a limited aperture of 50° in the vertical, and the azimuthal angles subtended at the receivers by the protuberance vary from 4.8° to 7.2°, i.e., a mean angle of $\varphi_c = 6.0°$. Here, we have implicitly assumed the source to be located at the origin of the azimuthal coordinate system. The salient reason for choosing a scatterer in a region NE relative to the apex of the arrays is that the plane of tilt of the long vertical array also lies in this direction. Therefore, the variation of the azimuthal angle at the sensors on the array will not be very significant. Also, the array is positioned such that it receives the scattered field in the backward direction at a range of about 1350 m from the discrete feature, as shown in Fig. 15.

The approach for estimating the scattered pressure from the reverberation data is similar to that outlined in Sec. III. In this case, however, we do not normalize the scattered pressure with respect to the ensonified area. Denoting the Fourier transform of the signal from the scatterer by $S(f)$, the directivity of the scatterer as a function of the polar (θ) angle is then estimated as

$$B(f, \theta_i, \varphi_c) = \frac{X_i(f, \theta_i, \varphi_c)}{S(f)}, \tag{13}$$

where X_i is the received signal at the ith sensor. The normalized scattering intensity (I) is then given by

$$I(f, \theta_i, \varphi_c) = 20 \log |B(f, \theta_i, \varphi_c)|, \tag{14}$$

where θ_i and φ_c are the angles subtended by the scatterer at the ith receiver on the long vertical array.

There is a caveat in this processing scheme. Equation (13) inherently assumes that all the energy arriving within this two-way travel time window at the receivers is from this feature only. However, due to the geometry of the scattering

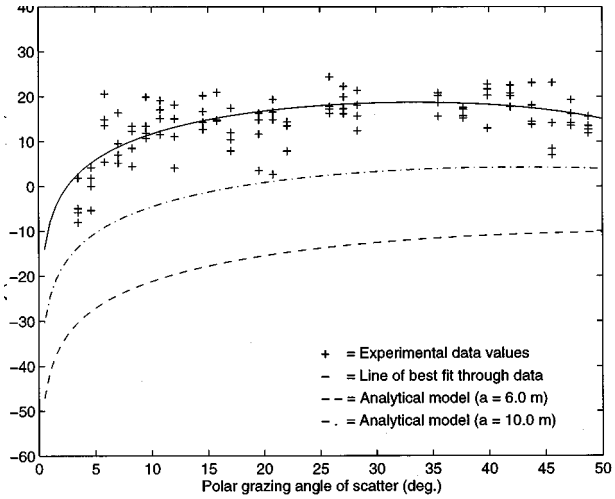


FIG. 16. Polar beampattern from experiments and analytical model for a frequency of $f=50$ Hz. The values from experiments are represented by (+), while the solid line (———) represents the best fit line through the data. The results from the analytical model are denoted by the dashed line (---) for $a=6.0$ m and by the dashed-dotted line (-·-) for $a=10.0$ m.

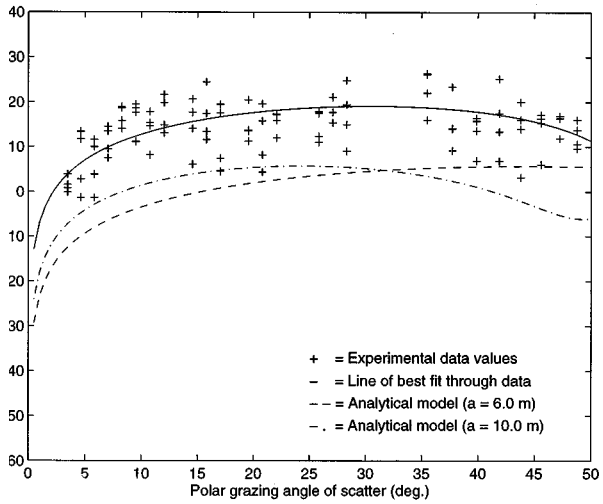


FIG. 17. Polar beampattern from experiments and analytical model for a frequency of $f=80$ Hz. The values from experiments are represented by (+), while the solid line (———) represents the best fit line through the data. The results from the analytical model are denoted by the dashed line (---) for $a=6.0$ m and by the dashed-dotted line (-·-) for $a=10.0$ m.

scenario, there will be multiple contributions from all such large features lying within the elliptical ensonified area. It is possible to reduce this ambiguity by using the beamformed signal instead of the source waveform directly. As the contributions from other scatterers will not add up coherently, we can expect to obtain a good estimate of the signal $S(f)$ from the feature, limited by the resolution of the array. However, it must be noted that $X_i(f)$ will still contain contributions from all scatterers within the ensonified area. We can obtain estimates of the distribution of scatterers within the ensonified area by assuming incoherent addition of energy. Letting M denote the effective number of such scatterers, the total scattered intensity from the whole ensonified area may be approximated as

$$L_{\text{total}}=I_{\text{single}}+10\log M. \tag{15}$$

We choose the material for the model to be elastic ice of density 910 kg/m^3, compressional wave speed of 3500 m s^{-1}, and shear wave speed of 1600 m s^{-1}. The ice plate is assumed to be 3 m thick and attached to the sphere of radius a along its equatorial plane. The coupled structure is completely submerged in water of density 1000 kg/m^3 and sound speed 1460 m s^{-1}. The point source is positioned at a distance of 1370 m and is incident at a grazing angle of 14.4°. We shall consider two frequency bands centered at 50 Hz and 80 Hz. Moreover, similar to the analysis in Sec. III, we shall average the results from experimental data over 10-Hz frequency bands.

Figure 16 shows the results from experimental data, and the analytical model for $f=50$ Hz. The experimental data values are denoted by the symbol ''+,'' and we have plotted the data from all four shots. Assuming the scattered field to consist of contributions from only the dipolar and quadrupolar terms, we model the data as

$$p_s=c_1\sin\theta+c_2\sin 2\theta,$$

where θ is the grazing angle of scatter, and c_1 and c_2 are the relative amplitudes of the polar components. The solid curve represents the least-squares best fit through the data points. We have also plotted the scattered field from our analytical realization of scattering from the approximate theoretical model for two values of the radius of the sphere, i.e., $a=6$ and 10 m. At a frequency of 50 Hz, these two curves correspond to ka values of 1.3 and 2.2, respectively. We observe reasonable qualitative agreement between the analytical and experimental results. Even though there is some qualitative agreement, there is discrepancy in magnitude. As stated earlier, this is due to contributions from multiple scatterers within the ensonified region in the experimental data. Also, the theoretical curves are essentially single frequency results while the experimental values have been averaged over 10-Hz bands. From Fig. 16, we observe that the difference in intensity levels between the experimental values and the analytical model is about 31 and 15 dB for the 6- and 10-m spheres, respectively. Using Eq. (15), we note that we can obtain a good match between the analytical results and experimental data by choosing $M=1260$, 32 for the 6- and 10-m spheres, respectively.

In Fig. 17, we have plotted the experimental values and the theoretical solutions for a higher frequency, i.e., $f=80$ Hz. In this case, the two theoretical curves for $a=6$ and 10 m correspond to ka values of 2.0 and 3.4, respectively. Once again we observe agreement in the trend and behavior between the analytical curves and the data with the better match being obtained with the 10-m sphere. The difference in intensity levels between the 10-m result and the experimental curve is about 14 dB corresponding to $M=25$. Analysis of two other frequency bands shows that a good match between experimental results due to multiple scatterers and that due to an isolated feature may be obtained

by adding 32 dB and 12 dB to the theoretical curves for 30 and 100 Hz, respectively.

V. SUMMARY

In this paper, we evaluated the three-dimensional acoustic scatter from features under the rough ice canopy in the Arctic Ocean using reverberation data collected on vertical and horizontal arrays. The analysis involved the isolation of discrete features followed by subsequent determination of their individual scattering characteristics. The analysis was carried out via a two-step matched-field processing algorithm. The first stage of the analysis, which involved near-field beamforming using the volumetric array, yielded estimates of scattering strengths for various frequency bands with the frequency regime (25–105 Hz) of interest in Arctic acoustics. It was demonstrated that scatter from ice protuberances is a frequency selective process. The beamformed map thus obtained was used to identify a single feature whose scattering pattern was subsequently evaluated. An idealized analytical model was also developed to compute acoustic scatter from a single idealized feature. Matching the experimental data which contains contributions from multiple scatterers with the analytical model yielded estimates of the distribution of protuberances under ice. In a future work, more exact statistics of distribution of ice features will be estimated by using the theoretical model in a statistical description.

ACKNOWLEDGMENTS

This work was supported by the Office of Naval Research (High Latitude Program). The authors are grateful to Tom Hayward of NRL for providing the reverberation data from the vertical line arrays.

[1] W. A. Kuperman and H. Schmidt, ''Rough surface elastic wave scattering in a horizontally stratified ocean,'' J. Acoust. Soc. Am. **79**, 1767–1777 (1986).

[2] W. A. Kuperman and H. Schmidt, ''Self-consistent perturbation approach to rough surface scattering in stratified elastic media,'' J. Acoust. Soc. Am. **86**, 1511–1522 (1989).

[3] K. LePage, ''Elastic scattering in oceanic waveguides,'' Ph.D. thesis, Massachusetts Institute of Technology, September 1992.

[4] L. M. Brekovskikh and Y. P. Lysanov, *Fundamentals of Ocean Acoustics* (Springer-Verlag, New York, 1990), pp. 182–215.

[5] V. Twersky, ''On scattering and reflection of sound by rough surfaces,'' J. Acoust. Soc. Am. **29**, 209–255 (1957).

[6] J. E. Burke and V. Twersky, ''On scattering and reflection by elliptically straited surfaces,'' J. Acoust. Soc. Am. **40**, 883–895 (1966).

[7] J. R. Fricke, ''Acoustic scattering from elastic ice: A finite difference solution,'' Ph.D. thesis, WHOI-MIT Joint Program, June 1991.

[8] P. Gerstoft and H. Schmidt, ''A boundary element approach to ocean seismoacoustic facet reverberation,'' J. Acoust. Soc. Am. **89**, 1629–1642 (1991).

[9] D. R. Burns and R. A. Stephens, ''Three-dimensional modeling of geoacoustic scattering from seafloor topography,'' J. Acoust. Soc. Am. **88**, 2338–2345 (1990).

[10] T. W. Dawson, ''Acoustic scattering in a three-dimensional oceanic waveguide using boundary integral equation methods,'' J. Acoust. Soc. Am. **90**, 2609–2622 (1991).

[11] T. J. Hayward and T. C. Yang, ''Low-frequency Arctic reverberation. I: Measurement of under-ice backscattering strengths from short-range direct returns,'' J. Acoust. Soc. Am. **93**, 2517–2523 (1993).

[12] T. Kapoor and H. Schmidt, ''Matched field estimation of scattering from Arctic ice—Analysis of the CEAREX 89 field experiments,'' J. Acoust. Soc. Am. **93**, 2419(A) (1993).

[13] A. B. Baggeroer, W. A. Kuperman, and H. Schmidt, ''Matched-field processing: Source localization in correlated noise as an optimum parameter estimation problem,'' J. Acoust. Soc. Am. **83**, 571–587 (1988).

[14] T. Kapoor and H. Schmidt, ''Acoustic scattering from a three-dimensional protuberance on a thin, infinite, submerged elastic plate,'' J. Acoust. Soc. Am. **102**, 256–265 (1997).

[15] T. K. Kapoor, ''Three-dimensional acoustic scattering from Arctic ice protuberances,'' Ph. D. thesis, MIT, June 1995.

[16] K. von der Heydt, N. R. Galbraith, A. B. Baggeroer, R. Muench, P. S. Guest, and K. L. Davidson, ''CEAREX A-Camp: Navigation, bathymetry, CTD, meteorology, and LOFAR data report,'' WHOI technical memorandum, 24 January 1991.

[17] Private communication with Tom Hayward, NRL, Washington DC, June 1993.

[18] N. R. Chapman, ''Source levels of shallow explosive charges,'' J. Acoust. Soc. Am. **84**, 697–702 (1988).

A novel approach to sound scattering by cylinders of finite length

Zhen Ye[a)]
Institute of Ocean Sciences, Sidney, British Columbia, Canada and Department of Physics, National Central University, Taiwan, Republic of China

(Received 5 March 1996; revised 14 July 1996; accepted 25 January 1997)

This paper presents a new method for studying bistatic scattering of acoustic waves by finite cylinders. The method is based on the assumption that a differential element of the cylinder scatters sound as though it were part of an infinite cylinder of the same radius. Under this assumption, the scattering function of cylinder body is derived rigorously using the Kirchhoff integral theorem. The effect of scattering by the ends of finite cylinders is also considered. It is shown that the result from the previous cylinder method [T. K. Stanton, J. Acoust. Soc. Am. **83**, 55–63 (1988)] is a special case of the present method. The new approach, particular-useful for calculating the bistatic scattering function, is compared to this previous method by various examples. The numerical computation shows that when the scattering direction is close to the incidence direction, the difference between the two methods is small, and the difference generally increases as the scattering direction moves increasingly away from the incidence direction. It is further shown that the difference between the two methods is more severe for weak scatterers. [S0001-4966(97)00508-0]

PACS numbers: 43.30.Gv, 43.20.Fn [JHM]

INTRODUCTION

The scattering of acoustic waves by submerged cylinders of finite length has attracted considerable interest from scientists over the past decades. This is not only because the finite length cylinder is a potential candidate for modeling many objects in underwater environments, such as plankton, fish, and sands, but it is also an essential building block for studies of acoustic scattering by more complicated elongated bodies.

Since no analytic solution can be derived to describe the scattering of sound by a finite cylinder, approximate methods are usually resorted to. To date, many useful approximate methods have been developed. These methods include the volume integral method,[1] when the density contrast effect is negligible, the T-matrix method,[2] the Kirchhoff approximate method,[3] the geometrical theory of diffraction (GTD),[4] the matched asymptotic expansions method,[5] the Sommerfeld–Watson transformation,[6] the boundary integral method,[7] and the cylinder method proposed by Junger[8] and by Stanton.[9] An evaluation of some of these approximate methods has been given by Partridge and Smith.[10] Brill and Gaunaurd[11] also compared GTD and the Kirchhoff approximate theory. A good review on sound scattering by various objects of finite size can be found in Gaunaurd and Werby.[2]

Among these approximate methods, the cylinder method by Stanton,[9] hereafter abbreviated as the cylinder method, is a very versatile model that can be used on a wide range of acoustic scattering problems, as pointed out by Partridge and Smith.[10] It can be applied to virtually any elongated object. However, there are a few concerns with the method. From the derivation,[9] it is clear that the scattering function is based on that of the broadside incidence on an infinite cylinder. The extension to an oblique incidence, leaving out the end effects, is done by heuristically introducing a phase factor [refer to Eq. (24) in Stanton[9]]. A justification of the phase factor is that the method can recover the well-known result when applied to a fixed rigid straight cylinder for the backward scattering.[12] Another concern is that, even for the broadside incidence, whether the heuristically added phase factor can ensure the validity for scattering at an arbitrary angle, i.e., bistatic scattering, is questionable. Moreover, the assumption that the volume flow per unit length of the scattered field of the finite cylinder is that of the infinite cylinder implies that the finite cylinder is regarded as a line scatterer. This is expected to be usually valid when the acoustic wavelength is much larger than the radius of the cylinder. Lastly, the contribution from each differential element of the cylinder is assumed to be independent of the angle between the scattering direction and the axis of the cylinder. A more detailed discussion on the cylinder method has been documented in Ref. 13. Hence, it seems imperative to improve the cylinder method so that these concerns can be removed.

The purpose of this article is to present an improved approach to the bistatic scattering of acoustic waves by finite cylinders, the scattering from cylinder bodies in particular. This method uses the modal-series solution for an infinite cylinder in the Kirchhoff integral theorem. The theorem can be found in, for instance, Hecht and Zajac.[14] Here we note that the Kirchhoff integral should not be confused with the Kirchhoff approximation. Because the finite cylinder is a geometry not analytically tractable, the calculation of the scattering function of a finite cylinder is necessarily based on certain assumptions. In this paper, we employ the assumption used in Ref. 8 that is, a differential length of the cylinder scatters sound as though it were part of an infinite cylinder of

a)Electronic mail: zhen@joule.phy.ncu.edu.tw

the same radius. The present method is considered rigorous under the approximation, thereby providing a more rigorous foundation to Stanton's approach, as we can show later that the Stanton formulation can be obtained as a special case of the present approach. Our focus in this article will be on sound scattering at arbitrary incidence and scattering angles.

Although the method to be presented is valid for sound scattering from finite cylinders with any material composition, we consider in this paper the fluid-filled cylinders for simplicity. In this regard, the present paper can be viewed as a conceptual layout of the improved method. It will become clear later that the extension to other cylinders is straightforward, and will be done in forthcoming papers.

I. THEORY

A. General consideration

We begin by considering the sound scattering by an arbitrary scatterer. We assume that the scatterer cannot support shear waves. The wave scattering from such a scatterer can be described by the Kirchhoff's scalar scattering theory.[14] The wave equation for the scattering wave is

$$(\nabla^2+k^2)P_s=0, \tag{1}$$

where ∇^2 is the Laplacian operator, k is the wave number ω/c with ω being the angular frequency, and c the sound speed in the medium. With the aid of Green's theorem, the scattering wave at a point $\mathbf{r}$ can be expressed in terms of the scattering wave and its gradient evaluated on the surface of the scatterer,

$$P_s(\mathbf{r})=-\frac{1}{4\pi}\int\int_S ds' \cdot\left[\frac{e^{ikr_1}}{r_1}\nabla_{r'}P_s(\mathbf{r}')-P_s(\mathbf{r}')\nabla_{r'}\left(\frac{e^{ikr_1}}{r_1}\right)\right], \tag{2}$$

where the integration is performed on the surface of the scatterer, $\mathbf{r}'$ is the coordinate of a point on the surface, $\mathbf{n}$ is an outwardly directed unit normal to the surface, and $r_1=|\mathbf{r}-\mathbf{r}'|$.

Known as the Kirchhoff integral theorem, Eq. (2) can be further simplified into [see Eq. (9.125) in Jackson[15]]

$$P_s(\mathbf{r})=-\frac{1}{4\pi}\int\int_S ds' \frac{e^{ikr_1}}{r_1}\mathbf{n}\cdot\left[\nabla_{r'}P_s(\mathbf{r})+ik\left(1+\frac{i}{kr_1}\right)\frac{\mathbf{r}_1}{r_1}P_s(\mathbf{r}')\right]. \tag{3}$$

At far field, i.e., $r\gg r'$, we approximate

$$r_1\approx r-\mathbf{r}'\cdot\frac{\mathbf{r}}{r}. \tag{4}$$

Under this condition, Eq. (3) becomes

$$P_s(\mathbf{r})=-\frac{1}{4\pi}\frac{e^{ikr}}{r}\int\int_S ds' e^{-ik\mathbf{r}\cdot\mathbf{r}/r}\mathbf{n}\cdot\left[\nabla_{r'}P_s(\mathbf{r})+ik\left(1+\frac{i}{kr}\right)\frac{\mathbf{r}}{r}P_s(\mathbf{r})\right]. \tag{5}$$

The scattering amplitude can be deduced from the radiation condition [see Eq. (9.124) in Ref. 15]:

$$P_s(\mathbf{r})\to f(\hat{\mathbf{r}},\hat{\mathbf{r}}_i)\frac{e^{ikr}}{r},\quad \text{as}\quad r\to\infty, \tag{6}$$

where $\hat{\mathbf{r}},\hat{\mathbf{r}}_i$ refer to the scattering and incidence directions, respectively. Thus the scattering function can be obtained from

$$f(\mathbf{r},\mathbf{r}_i)=-\frac{1}{4\pi}\int\int_S ds' e^{-ik\mathbf{r}\cdot\mathbf{r}/r}\mathbf{n}\cdot\left[\nabla_{r'}P_s(\mathbf{r}')+ik\frac{\mathbf{r}}{r}P_s(\mathbf{r}')\right], \tag{7}$$

where the higher-order term involving $1/r$ has been dropped, since we are concerned with the Sommerfeld radiation field here. In fact, the above formula is valid when $kr\gg 1$ and $r\gg R$, where R is the minimum radius of a sphere enclosing the scatterer (circumscribing sphere).

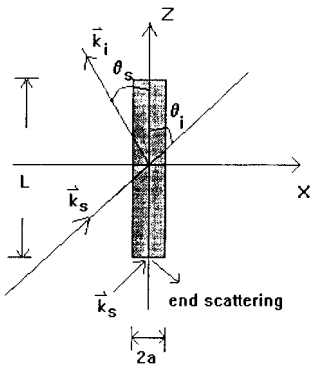
The above shows that the scattering wave at any spatial point can be evaluated by its values on the surface of the scatterer. In other words, once the values of the scattering wave on the surface of the scatterer are determined, the scattering wave at any point can be evaluated rigorously from Eq. (2). Unfortunately, except for a few simple geometries such as sphere, infinite cylinder, and prolate/oblate spheroid, no analytic solution can be deduced for the scattering wave at the surface of an arbitrary scatterer. Therefore, certain approximation has to be made for the scattering wave at the scattering surface in order to use the Kirchhoff integral theorem. A common approximation, known as the Kirchhoff approximation, is that the scattered wave is considered as an integral of the scattering from surface elements, each of which is assumed to have a reflection coefficient equal to an infinite plane surface.[3] Over the shadowed part of the scatterer, the scattering wave is taken as zero. The Kirchhoff approximation may suffer from some deficiencies, that is, it may not account for the material composition of targets, and is limited to scattering at high frequencies.[3]

In the following, we will propose a new approximation in the Kirchhoff integral theorem to study the scattering of acoustic waves by finite cylinders.

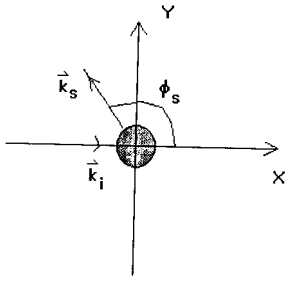
B. The formalism of the improved method

According to the above consideration, the essential task is to evaluate the scattering wave at the surface of the scatterer in order to apply the Kirchhoff integral theorem.

Consider a finite fluid cylinder. Following the assumption that a differential element of the cylinder scatters sound as though it were part of an infinite cylinder of the same radius,[8] we may approximate that the internal field, and also



(a) Side view



(b) Top view

FIG. 1. Geometry and coordinates for sound scattering from a finite cylinder.

the near field (scattering wave), is the same as that of the infinite long cylinder. This will allow us to use Eq. (2) to calculate the scattering function at far field.

The scattering of sound by an infinitely long cylinder is well known, and can be found in, for example, Morse[16] and Flax *et al.*[17] Here we simply list the result. Consider a unit incident plane wave in the $x-z$ plane. The axis of the cylinder is taken to be the z axis. In the cylindrical coordinates (Fig. 1), the incident wave can be written as

$$P_{\text{inc}}=e^{ikr\ \cos\ \phi\ \sin\ \theta_i+ikz\ \cos\ \theta_i}$$

$$=e^{ikz\ \cos\ \theta_i}\sum_{n=0}^{\infty} i^n\epsilon_n J_n(kr\ \sin\ \theta_i)\cos(n\phi), \tag{8}$$

where θ_i is the incident angle, ϕ is the azimuthal angle that sweeps through the plane perpendicular to the longitudinal cylinder axis, $J_n()$ is the nth order Bessel function of the first kind, and ϵ_n is the Neuman factor

$$\epsilon_n=\begin{cases}1, & n=0\\ 2, & n=1,2,3,\cdots\ .\end{cases}$$

The internal field can be solved as

$$P_{\text{int}}(\mathbf{r})=e^{ikz\ \cos\ \theta_i}\sum_{n=0}^{\infty} A_n J_n(k_1 r\ \sin\ \theta_i)\cos(n\phi), \tag{9}$$

where k_1 is the wave number inside the cylinder. The scattering wave can be written as

$$P_s(\mathbf{r})=e^{ikz\ \cos\ \theta_i}\sum_{n=0}^{\infty} B_n H_n^{(1)}(kr\ \sin\ \theta_i)\cos(n\phi), \tag{10}$$

in which $H_n^{(1)}()$ is the Hankel function of the first kind. The coefficients A_n, B_n are determined by the boundary conditions that the pressure and radial component of the particle velocity must be continuous at the boundary. We have

$$B_n=\frac{-i^n\epsilon_n[J_n(ka\ \sin\ \theta_i)-\beta_n J_n''(ka\ \sin\ \theta_i)]}{H_n^{(1)}(ka\ \sin\ \theta_i)-\beta_n H_n^{(1)'}(ka\ \sin\ \theta_i)}, \tag{11}$$

where

$$\beta_n=\frac{k\rho_1}{k_1\rho}\frac{J_n(k_1 a\ \sin\ \theta_i)}{J_n'(k_1 a\ \sin\ \theta_i)}=gh\frac{J_n(k_1 a\ \sin\ \theta_i)}{J_n'(k_1 a\ \sin\ \theta_i)},$$

where ρ_1, ρ are the mass density insider and outside the cylinder, respectively, a is the radius of the cylinder, $g=\rho_1/\rho$, $h=k/k_1$. In the above $J_n'(x)$ and $H_n^{(1)'}(x)$ denote $dJ_n(x)/dx$ and $dH_n^{(1)}(x)/dx$, respectively.

Under the assumption that a differential element of the cylinder scatters sound as though it were part of an infinite cylinder of the same radius, the scattering function of a finite cylinder can be obtained by substituting Eq. (10) into Eq. (7). The integration should be performed on the surface of the finite cylinder. (The contribution from the end of the finite cylinder is to be addressed later.) For convenience, we use the following convention: The incidence and scattering directions are expressed in terms of spherical coordinates while the points on the surface of the cylinder are expressed in the cylindrical coordinates. In this convention, the incidence direction is $(\theta_i,0)$, and the scattering direction is (θ_s,ϕ_s) in the spherical coordinates (Fig. 1). In this figure, $\mathbf{k}_i$ and $\mathbf{k}_s$ are the incident and scattering wave vectors and their norms equal k.

A computation leads to

$$f(\theta_s,\phi_s,\theta_i)=\frac{-iL}{\pi}\ \text{sinc}\ \Delta\sum_{n=0}^{\infty} B_n F_n(-i)^n\cos(n\phi_s), \tag{12}$$

in which

$$\text{sinc}\ \Delta=\frac{\sin(kL(\cos\ \theta_s-\cos\ \theta_i)/2)}{[kL(\cos\ \theta_s-\cos\ \theta_i)]/2}, \tag{13}$$

and

$$F_n=\frac{\pi}{2i}(ka)[H_n^{(1)'}(ka\ \sin\ \theta_i)J_n(ka\ \sin\ \theta_s)\sin\ \theta_i$$

$$-H_n^{(1)}(ka\ \sin\ \theta_i)J_n'(ka\ \sin\ \theta_s)\sin\ \theta_s]. \tag{14}$$

The major steps in the derivation of Eq. (12) are given in the Appendix.

From the derivation, we see that a scattering function from the cylinder body can be derived rigorously based on the assumption that a differential element of the cylinder scatters sound as though it were part of an infinite cylinder of the same radius.[8] The need for an added phase factor in Ref. 9 is removed. The dependence on the incidence and scattering angles shows up naturally from the derivation. We note that the present approach draws some similarities to the ap-

proach of Montroll and Hart[1] based on the volume integral. Using Green's theorem, one can show that when the density ratio between the medium and the scatterer is equal, the two approaches are expected to yield the same results.[18]

For comparison, we write the result from the cylinder method by Stanton as follows:[9]

$$f_{\mathrm{DCM}}(\theta_s,\phi_s,\theta_i)=\frac{-iL}{\pi}\ \mathrm{sinc}\ \Delta\sum_{n=0}^{\infty}B_n(-i)^n\cos(n\phi_s). \tag{15}$$

By the Wronskian identity

$$[J_n(x)H_n^{(1)'}(x)-J_n'(x)H_n^{(1)}(x)]=\frac{2i}{\pi x}, \tag{16}$$

it is easy to show the cylinder method in Eq. (15) is identical to the present result in Eq. (12) when $\theta_s=\theta_i$, or $\pi-\theta_i$.

To complete our discussion, the end effects of the finite cylinder need to be considered. We note here that the end effects may not be a major concern in many cases, such as when incidence and scattering are nearly normal to the axis, or when we consider an elongated deformed cylinder with small round ends. The latter case may be found in fish scattering. By recognizing that the ends are flat surfaces, the scattering from such a surface may be studied by the modified Kirchhoff integral which assumes, as aforementioned, that the scattered wave is considered as an integral of the scattering from surface elements, each of which is assumed to have a reflection coefficient equal to an infinite plane surface.[3] Over the shadowed part of the scatterer, the scattering wave is taken as zero. The scattering function can thus be written as

$$f_{\mathrm{end}}(\mathbf{r},\mathbf{r}_i)=\frac{i}{4\pi}\left[\frac{gh-1}{gh+1}\right]\int\int_S ds'\ e^{i(\mathbf{k}_i-\mathbf{k}_s)\cdot\mathbf{r}'}[\mathbf{k}_i-\mathbf{k}_s]\cdot\mathbf{n}, \tag{17}$$

which can be readily integrated as

$$f_{\mathrm{end}}(\theta_i,\theta_s,\phi_s)=\left[\frac{gh-1}{gh+1}\right]\left[\frac{-ika^2}{2}\right](\cos\theta_s-\cos\theta_i)e^{-ikL/2(\cos\theta_s-\cos\theta_i)}\frac{J_1(k'a)}{k'a}, \tag{18}$$

with

$$k'=k\left|\frac{\sin\theta_i-\sin\theta_s\cos\phi_s}{\cos\phi_0}\right|,$$

and

$$\phi_0=\tan^{-1}(\sin\theta_s\sin\phi_s/(\sin\theta_i-\sin\theta_s\cos\phi_s)).$$

In the above $J_1()$ is the first-order Bessel function of the first kind. In the cylinder models of Junger[8] and Stanton,[9] scattering by the ends is ignored. We note here that the present treatment of the scattering by the ends is not rigorous, and is limited perhaps to high-frequency scattering. In addition, in some situations, such as in short cylinders, multiple internal reflections may be also important. Nevertheless, the present treatment stemmed from the physical optics approach,[3] which may be the best compromise at this stage. Also, here we only considered the flat ends, because the major issue in this paper is to present an improvement of evaluation of the scattering from cylinder bodies.

The complete scattering function of a finite cylinder is given by the sum of the two results in Eq. (12) and Eq. (18). For convenience, we may, as usual, define a so-called target strength function

$$\mathrm{TS}=20\log_{10}|f(\theta_i,\theta_s,\phi_s)|. \tag{19}$$

II. NUMERICAL RESULTS AND DISCUSSION

In this section, we look at some numerical examples of finite cylinders. As shown in the above, the present method differs from the cylinder model of Stanton by the factors F_n in the modal series. These factors equal one for backscattering, forward scattering, and specular scattering. To probe the effects of these factors, we need to consider general bistatic scattering from the cylinders; that is, given the incidence direction we calculate the scattering function as a function of the scattering angle. Here the bistatic scattering means that the scattering direction is not in the opposite direction of the incidence. The importance of bistatic or multidirectional scattering of acoustic waves by a scatterer has now been increasingly recognized. For example, it has been shown that the forward scatter technique is useful for fish detection and quantification in shallow waters or river environments.[19] The current theory[19] assumes that only the scattering by a fish in the forward direction needs to be considered. This is only valid when the fish are distributed uniformly in the acoustic beam and the acoustic frequency is high. When these assumptions fail, a discrete model for fish detection in the forward scatter must be developed.[20] In this model, the bistatic scattering function of fish will be needed. In addition, apparently the bistatic scattering function of a fish is essential for using the new spatial filtering technique for locating fish in a river environment.[21]

Although it may be desirable to compare the present approach to many existing methods, we choose to compare mainly with Stanton's cylinder model.[9] This is not only because the present approach is mostly related to Stanton's method, and it has been shown that for some special cases the present result reduces to that of Stanton for backscattering, but also the cylinder model has been widely used in a number of important applications.[22] And, a comprehensive comparison between Stanton's cylinder method with other methods has been detailed by Partridge and Smith.[10] Their comparison can therefore be viewed as an indirect comparison between the present results and other results for backscattering. In addition, the present paper is focused on bistatic scattering by straight cylinders, and results from other methods are not readily available for comparison.

For convenience, throughout this article we use the following convention: The forward scattering means $\theta_s=\theta_i$, $\phi_s=0$ and the backscattering mean $\theta_s=\pi-\theta_i$, $\phi_s=\pi$; the backward direction scattering means $\phi_s=\pi$ and the forward direction scattering means $\phi_s=0$, while θ_s varies from 0 to π.

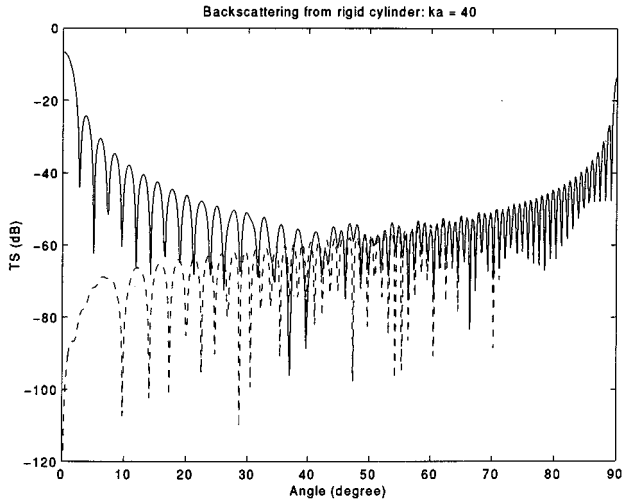


FIG. 2. Backscattering function in terms of target strength as a function of incidence angle for a rigid cylinder at $ka=40$: $L=0.12$ m, $a=0.0235$ m. The solid line refers to the total scattering, including contributions from the end and cylinder body. The dashed line is from the cylinder body only.

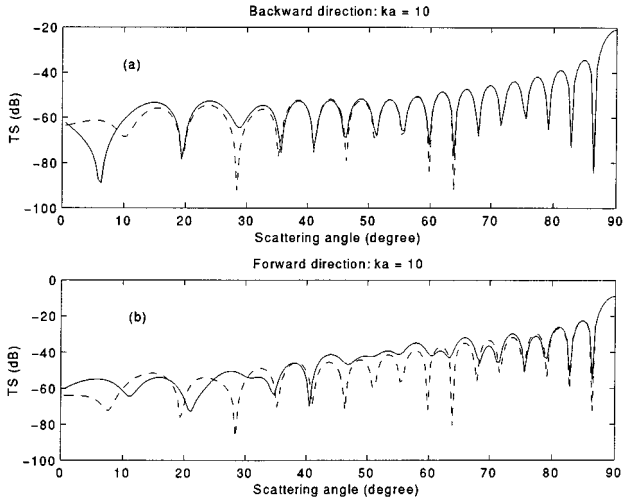


FIG. 3. Scattering function in terms of target strength as a function of scattering angle for a gas cylinder at $ka=10$: $L=0.1$ m, $a=0.01$ m, $g=0.001\,29$, $h=0.23$. The solid lines refer to the total scattering. The dashed lines are from the cylinder body only. The incidence angle $\theta_i=\pi/2$: (a) $\phi_s=\pi$; (b) $\phi_s=0$. In the plots, θ_s varies from 0 to 90 deg, and $\phi_s=\pi$ corresponds to the backward direction, whereas $\phi_s=0$ corresponds to the forward direction.

A. The end effects

Although it would be sufficient to only compare Eq. (12) and Eq. (15) in order to investigate the difference between the present theory and the existing cylinder model, we would like to first discuss the end effects. From Eq. (18), we see that the end effects may be negligible for weak scatterers, i.e., $g,h\sim 1$. However, for rigid ($g,h\gg 1$) or gas-filled ($g,h\ll 1$) cylinders, the scattering from the ends can be rather significant. In addition, the end effect increases as ka increases.

First we consider the backscattering ($\theta_s=\pi-\theta_i, \phi_s=\pi$) from a rigid cylinder: $L=0.12$ m, $a=0.0235$ m, and $ka=40$ (chosen to compare with the result in Ref. 10). In Fig. 2, we plot the scattering function in terms of target strength (abbreviated in the following as scattering function without confusion) as a function of incidence angle. The solid line refers to the scattering function, including the end and body contributions, while the dashed line refers to the body contribution. Here we see that the end effect becomes more and more important as the incidence angle decreases. Here we note that $\theta_i=0$ corresponds to the end-on incidence and $\theta_i=\pi/2$ refers to the broadside incidence. The end effect becomes noticeable when the incidence angle is greater than 40 deg from the broadside. GTD predicts that the end effect is significant when the incidence angle is greater than 25 deg from the broadside.[10]

The end effects can also be noticeable for bistatic scattering. For example, we consider scattering from a gas-filled cylinder: $g=0.001\,29$, $h=0.23$, $L=0.1$ m, $a=0.01$ m, and $ka=10$, and set the incidence at the broadside ($\theta_i=\pi/2$). In Fig. 3, the bistatic scattering is plotted as a function of the scattering angle. The solid line refers to the scattering function, including end and body contributions, while the dashed line refers to the body contribution. We see that the end effect becomes important when the scattering angle is greater than about 25 deg from the broadside in the backward case, while it is noticeable when the scattering angle is greater than about 5–10 deg from the broadside in the forward case.

The end effects can be negligible for weak scatterers (which could model zooplanktons). We consider: $g=1.04$, $h=1.04$, $L=0.1$ m, $a=0.01$ m, and $ka=10$. We set the incidence at the broadside ($\theta_i=\pi/2$). In Fig. 4, the bistatic scattering is plotted as a function of the scattering angle. The solid line refers to the scattering function, including end and body contributions, while the dashed line refers to the body

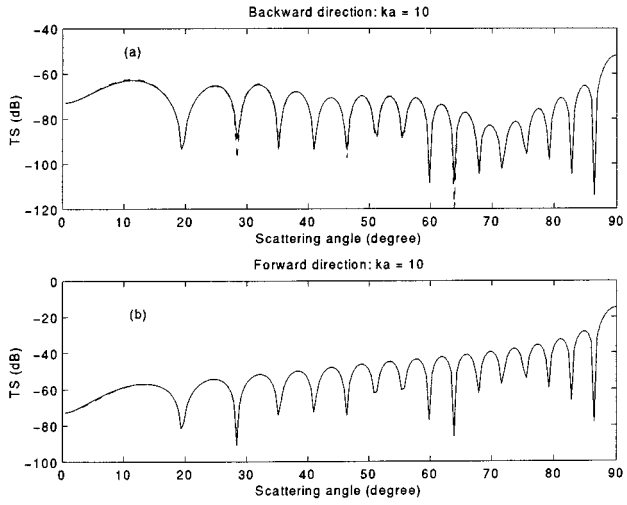


FIG. 4. Scattering function in terms of target strength as a function of scattering angle for a weak cylinder at $ka=10$: $L=0.1$ m, $a=0.01$ m, $g=h=1.04$. The solid lines refer to the total scattering The dashed lines are from the cylinder body only. The incidence angle $\theta_i=\pi/2$: (a) $\phi_s=\pi$; (b) $\phi_s=0$. In the plots, θ_s varies from 0 to 90 deg, and $\phi_s=\pi$ corresponds to the backward direction, whereas $\phi_s=0$ corresponds to the forward direction.

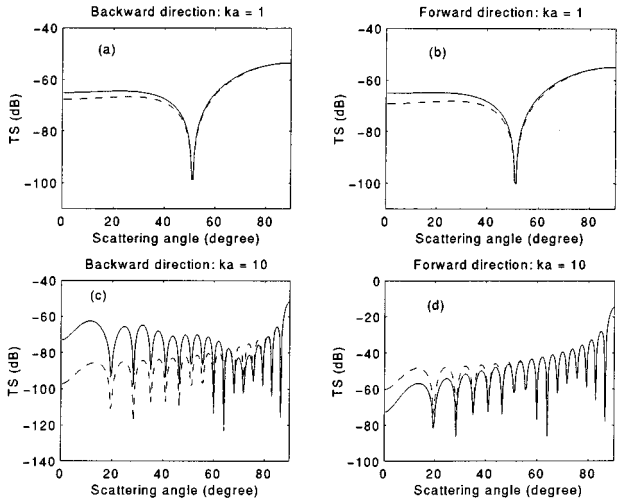


FIG. 5. Scattering function in terms of target strength as a function of scattering angle for a weak cylinder at $L=0.1$ m, $a=0.01$ m, $g=h=1.04$. The solid lines refer to the results from Eq. (12). The dashed lines are from Eq. (15). The incidence angle $\theta_i=\pi/2$: (a) $\phi_s=\pi$ and $ka=1$; (b) $\phi_s=0$ and $ka=1$; (c) $\phi_s=\pi$ and $ka=10$; (d) $\phi_s=0$ and $ka=10$. In the plots, θ_s varies from 0 to 90 deg, and $\phi_s=\pi$ corresponds to the backward direction, whereas $\phi_s=0$ corresponds to the forward direction.

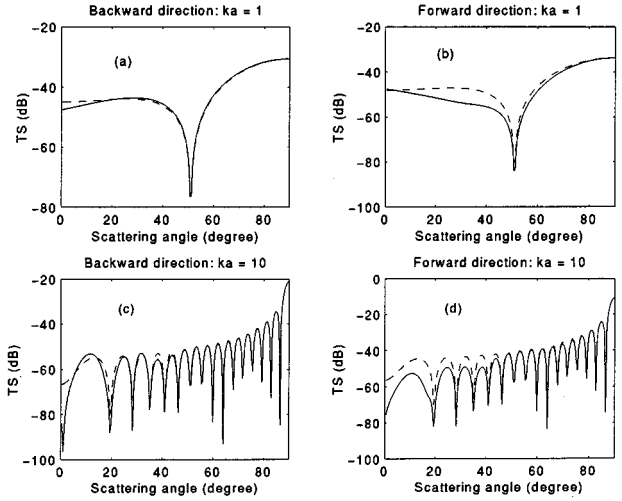


FIG. 7. Scattering function in terms of target strength as a function of scattering angle for a rigid cylinder: $L=0.1$ m, $a=0.01$ m. The solid lines refer to the results from Eq. (12). The dashed lines are from Eq. (15). The incidence angle $\theta_i=\pi/2$: (a) $\phi_s=\pi$ and $ka=1$; (b) $\phi_s=0$ and $ka=1$; (c) $\phi_s=\pi$ and $ka=10$; (d) $\phi_s=0$ and $ka=10$. In the plots, θ_s varies from 0 to 90 deg, and $\phi_s=\pi$ corresponds to the backward direction, whereas $\phi_s=0$ corresponds to the forward direction.

contribution. We see that the end effect is not important for the whole range of the scattering angles, as expected.

B. Comparison between the improved cylinder model and the previous cylinder model

We compare the results in Eq. (12) and Eq. (15). For our purpose, we consider the following size for the cylinder: $L=0.1$ m and $a=0.01$ m.

First we consider broadside incidence, i.e., $\theta_i=\pi/2$. Figures 5, 6, and 7 show the bistatic scattering function as a function of scattering angle for two different ka values: $ka=1$ and 10. In these figures, the solid lines refers to the computation from Eq. (12) and the dashed lines are for the results from the previous cylinder model in Eq. (15). In Fig. 5, we assume $g=h=1.04$ (weak scatterer). Figure 6 is for the gas cylinder, a potential candidate for modeling fish swimbladder.[22] The rigid cylinder case is presented in Fig. 7.

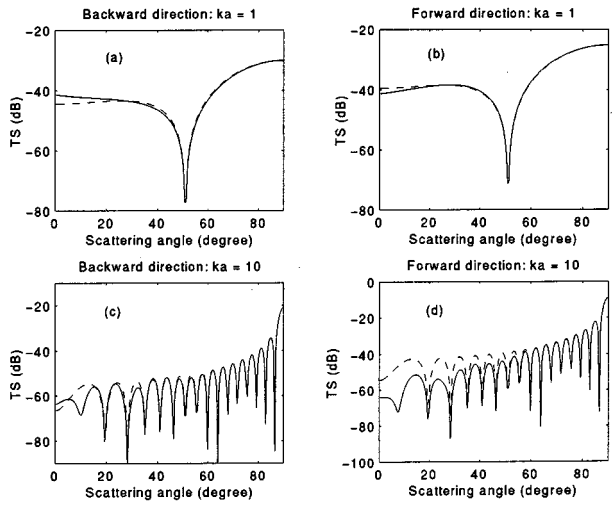


FIG. 6. Scattering function in terms of target strength as a function of scattering angle for a gas cylinder: $L=0.1$ m, $a=0.01$ m, $g=0.001\,29$, $h=0.23$. The solid lines refer to the results from Eq. (12). The dashed lines are from Eq. (15). The incidence angle $\theta_i=\pi/2$: (a) $\phi_s=\pi$ and $ka=1$; (b) $\phi_s=0$ and $ka=1$; (c) $\phi_s=\pi$ and $ka=10$; (d) $\phi_s=0$ and $ka=10$. In the plots, θ_s varies from 0 to 90 deg, and $\phi_s=\pi$ corresponds to the backward direction, whereas $\phi_s=0$ corresponds to the forward direction.

From these figures we see that; (1) the two models agree well with each other when the scattering angle is close to the incidence angle, i. e., $\theta_s\sim\pi/2$; (2) the discrepancies between the two models are more and more noticeable when the scattering direction moves away from the incidence direction; (3) for the weak scatterer case ($g=h=1.04$) the difference between the two methods can be rather significant, even as the scattering angle differs from the incidence angle by merely 15 deg in the backward direction [Fig. 5(c)]; (4) in the gas cylinder case, the difference between the two methods is more pronounced in the forward direction scattering for the larger ka [Fig. 6(d)]; (5) in the rigid cylinder case, the difference is significant for both large and small ka in the forward direction [Fig. 7(b) and (d)]; (6) the difference is usually small for the smaller ka cases, except for the forward direction scattering by the rigid scatterer [Fig. 7(b)]; (7) clearly, the difference between the two models arises from the additional factors F_n in Eq. (12), which are a result of the proper account of the contributions from the differential elements of the cylinders.

Taking into account the end effects, the difference between the two methods may still be significant. In Fig. 8, we plot the scattering function as a function of scattering angle for the rigid cylinder at $ka=10$, incorporating the scattering from the ends. In this figure, the solid lines are the results from the present theory in Eqs. (12) and (18), while the dashed lines are from Eqs. (15) and (18). Here we see that compared with Fig. 7(c) and (d), the end effect becomes

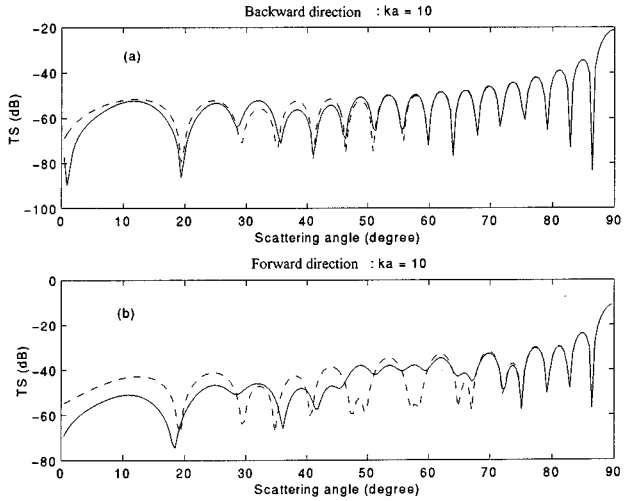


FIG. 8. Comparison between the previous cylinder model and the present model, including the end effects. The scattering function in terms of target strength is plotted as a function of scattering angle for a rigid cylinder: $L=0.1$ m, $a=0.01$ m. The solid lines refer to the results from Eqs. (12) and (18). The dashed lines are from Eqs. (15) and (18). The incidence angle $\theta_i=\pi/2$: (a) $\phi_s=\pi$ and $ka=10$; (b) $\phi_s=0$ and $ka=10$. In the plots, θ_s varies from 0 to 90 deg, and $\phi_s=\pi$ corresponds to the backward direction, whereas $\phi_s=0$ corresponds to the forward direction.

important as the scattering angle differs from the incidence angle by about 35–40 deg in the backward direction scattering and by about 25 deg in the forward direction scattering, due to the interference between the body and end contributions.

We may also consider oblique scattering. As an example, Fig. 9 presents such a case for the weak cylinder. The incidence angle is set at 70 deg. The differences between the two models as a function of scattering angle are plotted in Fig. 9(b) and (d), respectively. Here we see that the differences for the method are small for near broadside scattering. The significant difference does appear when the scattering angle is away from the broadside by more than 40 deg, as shown in Fig. 9(b) and (d).

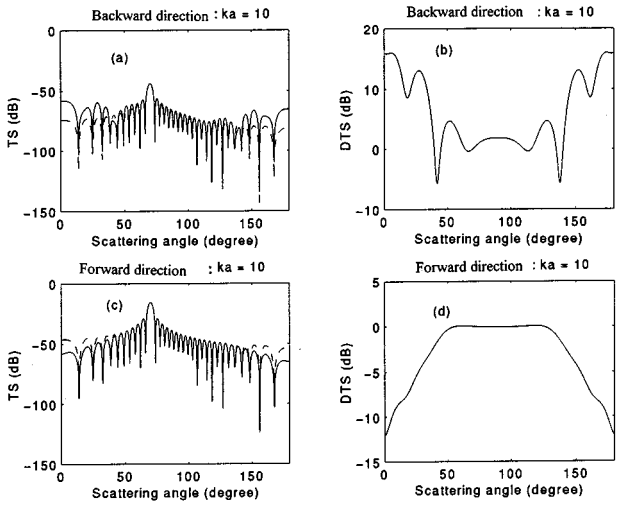


FIG. 9. Scattering function in terms of target strength as a function of scattering angle for a weak cylinder: $L=0.1$ m, $a=0.01$ m, $g=h=1.04$. The solid lines refer to the results from Eq. (12). The dashed lines are from Eq. (15). The incidence angle $\theta_i=70$ deg: (a) $\phi_s=\pi$ and $ka=10$; (b) the difference between the solid and dashed lines in (a); (c) $\phi_s=0$ and $ka=10$; (d) the difference between the solid and dashed lines in (c). In the plots, θ_s varies from 0 to 90 deg, and $\phi_s=\pi$ corresponds to the backward direction, whereas $\phi_s=0$ corresponds to the forward direction.

III. SUMMARY

In this paper, we presented a novel method for calculating sound scattering function by finite cylinders. The present method is based on the previous assumption that a differential element of the cylinder scatters sound as though it were part of an infinite cylinder of the same radius. Under this assumption, the scattering function is derived rigorously using the Kirchhoff integral theorem. It was shown that the present method can naturally yield the cylinder method proposed by Stanton for backscattering, forward scattering, and specular scattering, thereby providing a footage to Stanton's model. In this article we limited our discussion to cases when the incidence, scattering, and the axis of cylinders are in the same plane. The effect of scattering by the ends of finite cylinders has also been considered. It was found that the present model differs from that of Stanton by some factors in the modal-series solution.

The present model has been compared to the previous cylinder model via various examples. The numerical computation has shown that when the scattering direction is close to the incidence direction, the difference between the two methods is small, and the difference generally increases as the scattering direction moves increasingly away from the incidence direction. It was further shown that the difference between the two methods is more severe for weak scatterers. For the backscattering or straightforward scattering, the previous cylinder model may be sufficient. When the arbitrary bistatic scattering is of concern, the present model seems to be a more appropriate candidate.

Although only fluid cylinders are considered in this paper, it is clear from the derivation that the present method can be easily extended to study sound scattering by any other cylinders, or elastic shells.[2,23] The extension may be done by substituting the appropriate solution for the scattering wave at the outer surface into the Kirchhoff integral.

ACKNOWLEDGMENTS

The author thanks David M. Farmer and Masahiko Furusawa for valuable discussion and suggestions. The comments from C. S. Clay and D. Chu are also greatly appreciated. The referees are thanked for many useful comments and constructive suggestions. The work is supported by the National Program for Fisheries Acoustics, Department of Fisheries and Oceans, Canada, and the College of Science, National Central University.

APPENDIX

In this Appendix we provide a proof of Eq. (12).

As mentioned earlier, we use the following convention: The incidence and scattering are expressed in terms of spherical coordinates, while the points on the surface of the cylinder are expressed in the cylindrical coordinates. In this

convention, the incidence direction is $(\theta_i,0)$, and the scattering direction is (θ_s,ϕ_s) in the spherical coordinates. We also note that the solution for an infinite cylinder is expressed in terms of the cylindrical coordinates. Therefore we have

$$\mathbf{r}=r(\sin\theta_s\cos\phi_s,\ \sin\theta_s\sin\phi_s,\ \cos\theta_s),$$

$$\mathbf{r}=a(\cos\phi,\ \sin\phi,z),$$

and

$$\mathbf{n}=(\cos\phi,\ \sin\phi,0).$$

Using Eq. (4), these leads to

$$|\mathbf{r}-\mathbf{r}'|\approx r-a\sin\theta_s\cos\phi_s\cos\phi - a\sin\theta_s\sin\phi_s\sin\phi-z\cos\theta_s. \tag{A1}$$

The scattering function can be obtained from

$$f(\theta_s,\phi_s,\theta_i)=-\frac{1}{4\pi}\int_{-L/2}^{L/2}dz\int_0^{2\pi}d\phi\ a e^{-ika\sin\theta_s\cos\phi_s\cos\phi-ika\sin\theta_s\sin\phi_s\sin\phi-ikz\cos\phi_s}\left[\left.\frac{\partial P_s(r',z,\phi)}{\partial r'}\right|_{r'=a} + ik\mathbf{n}\cdot\frac{\mathbf{r}}{r}P_s(a,z,\phi)\right]. \tag{A2}$$

This equation can reduce to Eq. (12) by using the following relations:

$$\int_0^{2\pi}d\phi\ \cos(n\phi)e^{-ika(\sin\theta_s\cos\phi_s\cos\phi+\sin\theta_s\sin\phi_s\sin\phi)} = 2\pi(-i)^nJ_n(ka\sin\theta_s)\cos(n\phi_s), \tag{A3}$$

$$\int_0^{2\pi}d\phi\ \cos(n\phi)e^{-ika(\sin\theta_s\cos\phi_s\cos\phi+\sin\theta_s\sin\phi_s\sin\phi)}\times(\sin\theta_s\cos\phi_s\cos\phi+\sin\theta_s\sin\phi_s\sin\phi) = 2\pi(-i)^n(i\sin\theta_s)J_n'(ka\sin\theta_s)\cos(n\phi_s). \tag{A4}$$

The scattering wave at the boundary is approximated by that of the infinite cylinder,

$$P_s(a,z,\phi)=e^{ikz\cos\theta_i}\sum_{n=0}^{\infty}B_nH_n^{(1)}(ka\sin\theta_i)\cos(n\phi), \tag{A5}$$

and its derivative thus by

$$\left.\frac{\partial P_s(r',z,\phi)}{\partial r'}\right|_{r'=a} = e^{ikz\cos\theta_i}\sum_{n=0}^{\infty}(k\sin\theta_i)B_nH_n^{(1)'}(ka\sin\theta_i)\cos(n\phi), \tag{A6}$$

$$\int_{-L/2}^{L/2}dz\ e^{ikz\cos\theta_i-ikz\cos\theta_s} = L\frac{\sin(kL(\cos\theta_i-\cos\theta_s)/2)}{kL(\cos\theta_i-\cos\theta_s)/2}, \tag{A7}$$

and

$$\mathbf{n}\cdot\mathbf{r}/r=\sin\theta_s\cos\phi_s\cos\phi+\sin\theta_s\sin\phi_s\sin\phi. \tag{A8}$$

[1] E. W. Montroll and R. W. Hart, ''Scattering of plane waves by soft obstacles. II. Scattering by cylinders, spheroids, and disks,'' J. Appl. Phys. **22**, 1278–1289 (1951).

[2] G. C. Gaunaurd and W. F. Werby, ''Acoustic resonance scattering by submerged elastic shells,'' Appl. Mech. Rev. **43**, 171–208 (1990).

[3] G. C. Gaunaurd, ''Sonar cross sections of bodies partially insonified by finite sound beams,'' IEEE J. Ocean Eng. **OE-10**, 213–230 (1985).

[4] J. B. Keller and D. S. Ahluwalia, ''Diffraction by a curved wire,'' SIAM (Soc. Ind. Appl. Math.) J. Appl. Math. **20**, 390–405 (1971).

[5] M. Tran Van Nhieu and F. Ywanne, ''Sound scattering by slender bodies of arbitrary shape,'' J. Acoust. Soc. Am. **95**, 1726–1733 (1994).

[6] M. L. Rumerman, ''Application of the Sommerfeld–Watson transformation to scattering of acoustic waves obliquely incident upon cylindrical shells,'' J. Acoust. Soc. Am. **91**, 2502–2509 (1992).

[7] H. A. Schenck, ''Improved integral formation for acoustic radiation problems,'' J. Acoust. Soc. Am. **44**, 41–58 (1968).

[8] M. C. Junger, ''Scattering by slender bodies of revolution,'' J. Acoust. Soc. Am. **72**, 1954–1956 (1982).

[9] T. K. Stanton, ''Sound scattering by cylinders of finite length. I. Fluid cylinders,'' J. Acoust. Soc. Am. **83**, 55–63 (1988).

[10] C. Partridge and E. R. Smith, ''Acoustic scattering from bodies: Range of validity of the deformed cylinder method,'' J. Acoust. Soc. Am. **97**, 784–795 (1995).

[11] D. Brill and G. C. Gaunaurd, ''Approximate descriptions of the sound fields scattered by insonified waves obliquely incident upon cylindrical shells,'' J. Acoust. Soc. Am. **93**, 71–79 (1993).

[12] R. J. Urick, *Principles of Underwater Sound* (McGraw-Hill, New York, 1983).

[13] Z. Ye and M. Furusawa, ''Modeling of target strength of swimbladder fish at high frequencies,'' J. Acoust. Soc. Jpn. (E) **16**, 371–379 (1995).

[14] E. Hecht and A. Zajac, *Optics* (Addison-Wesley, London, 1974).

[15] J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1975).

[16] P. M. Morse, *Vibration and Sound* (AIP, New York, 1983).

[17] L. Flax, V. K. Varadan, and V. V. Varadan, ''Scattering of an obliquely incident acoustic wave by an infinite cylinder,'' J. Acoust. Soc. Am. **68**, 1832–1835 (1980).

[18] D. S. Jones, *Acoustic and Electromagnetic Waves* (Clarendon, Oxford, 1986), pp. 478–481.

[19] Z. Ye, T. Curran, and D. Lemon, ''Fish detection by acoustic scintillation technique,'' ICES J. Mar. Sci. **53**, 317–322 (1996).

[20] A. Ishimaru, *Wave Propagation and Scattering in Random Media* (Academic, New York, 1978), Chap. 14.

[21] David Farmer (private communication).

[22] C. S. Clay, ''Composite ray-mode approximations for backscattered sound from gas-filled cylinders and swimbladders,'' J. Acoust. Soc. Am. **92**, 2173–2180 (1992).

[23] A. Harari and B. E. Sandman, ''Radiation and vibrational properties of submerged stiffened cylindrical shells,'' J. Acoust. Soc. Am. **88**, 1817–1830 (1990).

On the use of matched-field processing and hybrid algorithms for vertical slice tomography

Michael I. Taroudakis
Department of Mathematics, University of Crete, 714 09 Heraklion, Crete, Greece and Institute of Applied and Computational Mathematics, Foundation for Research and Technology-Hellas, P.O. Box 1527, 711 10 Heraklion, Crete, Greece

Maria G. Markaki
Institute of Applied and Computational Mathematics, Foundation for Research and Technology-Hellas, P.O. Box 1527, 711 10 Heraklion, Crete, Greece

(Received 29 December 1996; accepted for publication 25 January 1997)

Matched-field processing and a hybrid scheme for vertical slice tomography are studied in characteristic cases of range-independent and range-dependent ocean environments using simulated data obtained at discrete points in the sound field along a vertical line, representing a vertical array of hydrophones. The matched-field processing is associated with a Genetic Algorithm, while the basic characteristic of the hybrid scheme is the application of matched-field processing for the definition of a reference environment to be subsequently used in connection with a modal-phase inversion scheme. In all cases studied, the sound-speed profiles are reconstructed with the use of empirical orthogonal functions. It is shown that the reconstruction of either the single sound-speed profile characterizing the range-independent environment, or the set of profiles characterizing an eddy in the range-dependent environment is achieved with good accuracy, despite the fact that a very wide search space has been considered. The apparent nonuniqueness in the solution of the inverse problem for the range-dependent case is remedied by applying physical constraints in the structure of the sea eddy and averaging over the admissible solutions obtained by the matched-field inversion scheme. [S0001-4966(97)04607-9]

PACS numbers: 43.30.Pc, 43.60.Gk [JHM]

INTRODUCTION

Monitoring the sea environment by acoustic means is a practice continuously gaining interest among researchers working in the fields of marine technology, oceanography, and environmental engineering. Since the introduction of ocean acoustic tomography by Munk and Wunsch, who based the data inversion procedure on the measurements of the ray travel time of known acoustic signals,[1] a number of different approaches have been proposed, classified mainly according to the characteristics of the acoustic measurements they utilize, in connection with the inversion methods applied to the acoustic data. Among them, modal travel time[2] and peak arrival methods[3] use information on the arrival pattern of the signal obtained in the time domain, whereas the modal-phase approach uses modal information in the frequency domain.[4–6] The deployment of an array of hydrophones to be used as the measuring device makes it feasible to apply matched-field techniques for source localization and general environmental parameter estimation, and this approach has been extensively studied over the last years.[7–11] Since ocean acoustic tomography is basically referred to the recovery of properties of the water column, all the associated inversion schemes lead to the estimation of the sound velocity structure in the water column, but simultaneous recovery of bottom geoacoustic parameters and source location using matched-field processing is also possible.

A main disadvantage of the methods developed so far, based on travel time (ray, mode, peak) or modal-phase measurements, is that they assume good *a priori* knowledge of the environment under reconstruction, whereas matched-field tomography can be applied in more general cases where there is limited information on the environment, since it is not necessarily based on the determination of a reference environment. The main characteristics of matched-field inversion schemes are that they are formulated conventionally through associated optimization problems for the estimation of the set of parameter values that maximizes an appropriate objective function, which is determined to represent the mismatch between the observed and computed data.

Since repeated changes of the input parameters of some suitable direct model and the solution of the associated problem are fundamental in matched-field processing schemes, they might be proven very slow and, accordingly, the role of the algorithm used to control the change in such a way that an acceptable solution is reached within a reasonable time avoiding local optima traps is essential.

The main objective of this paper is to study the performance of matched-field tomography in connection with a genetic algorithm serving as the iterations guiding tool. The matched-field approach supported by the genetic algorithm was tested in three characteristic cases chosen so that both shallow and deep water environments as well as range-dependent cases are considered.

Additionally, matched-field processing is also used in the framework of a new hybrid scheme employing modal-phase technique. This scheme is particularly suitable in cases when accuracy of the recovered structure is demanded even

when no reference environment has been determined.

The deep water case is referred to the western Mediterranean basin, where a tomography experiment with single source–receiver pairs[3] has been performed. It is shown here that the average sound-speed profile along a specific propagation path can be effectively recovered using matched-field technique, provided that the receiving device which was used in the actual experiment is replaced by a vertical array of hydrophones.

The range-dependent case is referred to a shallow water environment, with a cold eddy. This eddy has been used as the test case for the demonstration of the applicability of modal inversion techniques to recover range-dependent structures in the sea environment at a vertical slice.[6,12] It is shown that matched-field tomography, in association with a very wide-search space, gives acceptable results without the assumption of good *a priori* knowledge of a reference environment.

The hybrid scheme has been applied in the range-dependent case. It is shown that the modal-phase inversion scheme can be used complimentary to the matched-field processing algorithm to improve substantially its results.

The structure of the paper is as follows: Section I presents the general framework of the matched-field processing for underwater acoustics and especially for ocean acoustic tomography. Also the genetic algorithms are referred to as an effective tool for the systematic change of the input parameters and to the special development made in the present work. Section II explains briefly the problem of range dependence of the recoverable parameters, in connection with the adopted inversion scheme. Section III presents the modal-phase approach to be used in the hybrid inversion scheme, which is described in Sec. IV. Section V is devoted to the presentation of the test cases studied, and the results obtained by the two approaches. Special reference is made to the specific genetic algorithms incorporated in the matched-field processing for each case. The conclusions of the study are presented in Sec. VI. The Appendix contains a short description of the procedure followed for the application of deterministic crowding with threshold accepting, which is a subroutine needed for the efficient application of the genetic algorithm in test case III.

I. MATCHED-FIELD PROCESSING AND GENETIC ALGORITHMS

Matched-field processing (MFP) has been recently applied to underwater acoustics for solving general classes of inverse problems. It makes use of the spatial characteristics of the acoustic field in the water column for source localization and for the estimation of the environmental parameters, such as bottom structure and water column properties. The principles of matched-field processing for underwater acoustics can be found in the book by Tolstoy,[13] while comments on the problems encountered when applying MFP in inverse problems of this area, can be found in Ref. 14. The greater part of the present work deals with the application of matched-field processing in a typical problem of ocean acoustic tomography, trying to define an optimum search for the determination of the ''best solution'' to the problem.

A. The problem of ocean tomography

The acoustic field in the water column depends on the sound-speed structure, as well as on the boundary geometry and the bottom properties. Recovery of the sound speed as a function of space is the starting point of oceanographers for inferring oceanographic features in the water column. Neglecting time dependence, the sound speed $c(\mathbf{x})$ is generally a function of all spatial coordinates. In a cylindrical coordinate system, c is a function of r, z, and ϕ, $(c(r,z,\phi))$. For simplifying the calculations, the general 3-D problem is reduced to a 2-D one by assuming axial symmetry and thus an acoustic field independent of ϕ. The problem is now defined at a plane vertical to the sea surface, and the whole procedure is referred as ''vertical slice tomography.'' This will be the approach adopted in this paper.

All methods dealing with acoustic tomography are associated with the calculation of the channel transfer function, which is simply the solution of the acoustic wave equation, reduced for a point harmonic source:

$$\nabla^2 p(r,z)+k^2(r,z)p(r,z)=-\frac{1}{2\pi r}\,\delta(r)\,\delta(z-z_0) \qquad (1)$$

under the appropriate boundary conditions and a Sommerfield radiation condition for the behavior of the field at infinity. Here, $k(r,z)$ is the wave number $[k(r,z)=\omega/c(r,z)$, $\omega=2\pi f$, f being the source frequency]. As it will become clear in the sequel, we will adopt here the normal-mode representation for defining its solution.

The inverse problem of ocean acoustic tomography that will be treated in the present work is defined as follows:

> Given a vector of measurements of the complex acoustic field at N discrete points in the water column (in a vertical line), estimate the sound-speed function $c(r,z)$ corresponding to this set of measured data.

This problem falls within the general class of discrete inverse problems.[15] The set of data is given as a vector of elements corresponding to discrete measurements of the acoustic pressure and the recoverable parameters are generally given at discrete points in the water column.

The problem is of course ill-posed. Despite the fact that there is no theory in support of the existence and uniqueness of a solution to this problem under the specifications described above, computational efforts based on well-posed direct problems have shown that the calculation of the solution to the inverse problem is plausible in many cases.

B. The processor

Recent results derived by application of matched-field processing for ocean acoustic tomography have shown that the recovery of the sound speed as a function of range and depth is possible, provided that the search algorithm is associated with an effective processor such as the Bartlett processor described below. This processor is used in this work to crosscorrelate the measured with the modeled fields. It is defined as follows:

Consider an array of hydrophones consisting of N elements. Let $\mathbf{F}$ be the vector of the complex pressure field data

in these hydrophones and $\mathbf{w}(\mathbf{c})$ be a normalized prediction of the acoustic field on the array, due to some ''test'' sound-speed structure denoted by $\mathbf{c}$. In simple words, the elements of the vector $\mathbf{w}$ are complex pressures calculated at the field points corresponding to the hydrophone locations, by solving the direct acoustic propagation problem defined above.

The Bartlett processor is defined as

$$P(\mathbf{c}) = \mathbf{w}^{+}\mathbf{C}\mathbf{w}, \tag{2}$$

where $\mathbf{w}^{+}$ is the conjugate transpose of $\mathbf{w}$ and $\mathbf{C} = \langle \mathbf{F}\mathbf{F}^{+} \rangle$ is the cross-spectral data matrix. For N array elements, it is a $N \times N$ matrix representing an average of cross-spectral matrices, $\mathbf{F}\mathbf{F}^{+}$, formed from the complex pressure field data recorded several times within a time window (the brackets denote averaging). In real world the set of data includes noise due to both physical and manmade sources.

In order to simulate a real tomography experiment, the noise should be included in the ''measured'' field, which is in fact the field obtained by applying the forward propagation model with the parameters of the actual environment.

This is the case of the present work. In the simulated real ''data'' considered here, noise has been added, treated as Gaussian, white and uncorrelated one, and enters the calculations by the addition of the quantity $N_0 = 10^{-\mathrm{SNR}/10}$ to the diagonal of the estimated cross-spectral matrix $\mathbf{C}$. SNR denotes the average input signal-to-noise ratio. That is, for our simulations

$$\mathbf{C} = \mathbf{S}\mathbf{S}^{+} + N_0 \mathbf{I}_N, \tag{3}$$

where $\mathbf{S}$ is the simulated signal calculated at the array elements (always consisting of N hydrophones) and $\mathbf{I}_N$ is the $N \times N$ identity matrix.[13]

Using a normal-mode approach for studying acoustic propagation and determining the solution of the direct problem, we note that in order to match the field data, the MFP should in principle use all the modes allowed by the waveguide, even if their number, M, is greater than the number of hydrophones, N. In practice, when the signal is dominated by a small number of modes, $M' < N$, including the Nth and higher-order modes, makes little difference.[16] Moreover, if only a limited number of modes are used in the replica field, the result can be considered as equivalent to filtering the remaining modes out of the data (using a particular mode filtering algorithm). While this process no longer corresponds to the ''best'' match in the hydrophone space, it can be shown to lead again to an effective suppression of the sidelobes,[17] and a considerable reduction of the computation time, as confirmed in our second test case.

C. The genetic algorithms

The search technique incorporated in the MFP is a critical factor determining the performance of the scheme, which is defined by the speed of the calculations and the reliability of the solutions provided.

Genetic algorithms (GAs) have been used in inverse problems of underwater acoustics by Gerstoft and other researchers[18–21] with encouraging results. They exhibit impressive efficiency in optimizations, involving discontinuous, noisy, highly dimensional and multimodal objective functions,[22] as is the case with most matched-field processors. The GAs will be used here as search schemes.

A simple GA, as defined by Goldberg,[22] proceeds in two stages: It starts with some population of samples chosen from the search space. Then, selection is applied to create an intermediate population of above-average members. Further on, crossover (that is, recombination) and mutation (random bit change) are applied to the intermediate population to create the next one. The process of going from the current population to the next constitutes one generation in the execution of a GA.

The evaluation function must be relatively fast, since a GA works with a population of potential solutions, and incurs the cost of evaluating this population. The relative simplicity of the Bartlett processor was thus the main reason for its choice. We should point here that the normal-mode model used for the range-independent cases 1 and 2, SNAP,[23] can be parallelized for modes computed,[24] reducing the evaluation time for each member of the population. GAs, too, are intrinsically parallel; their implementation on parallel architectures that has recently received increased attention could reduce the total evaluation time according to the population size and the number of processors used.

Besides the evaluation (object) function, the other problem-dependent component of a genetic algorithm is the encoding of parameters to be recovered. Here, we use the concatenated, multiparameter, mapped binary coding. Each variable is mapped linearly from a specified interval $[U_{\min}, U_{\max}]$ to $[0, 2^l]$, and thus it is coded as a l-bit substring. The precision of the mapped coding is:

$$\pi = \frac{U_{\max} - U_{\min}}{2^l - 1},$$

and it is assumed that the resolution provided is ''visible'' at the output. A very fine precision enlarges the search space too much, while a coarse one may lose peaks of the fitness landscape. Thus it was determined empirically as a compromise between these extrema. It is finer for the first coefficient as one would expect, since it influences most the sound-speed profiles.

Although the GAs rapidly locate the region in which a global optimum exists, they do not locate the optimum with similar speed. A GA must be provided with a payoff incentive to exploit small differentials; that is, between the main peak and the sidelobes in the fitness landscape. A typical solution involves providing feedback to the GA in the form of a dynamically scaled fitness function in order to maintain sufficient selective pressure between competing individuals in the current population.[22] Unfortunately, there is currently no theory to support a suggestion on how to achieve an optimal scaling strategy. Extensive experimentation in the case of the range-independent environment of test case 1 gave excellent results. The same algorithm gave equally good results for the deep water environment of test case 2, for similar values of source frequency and range.

But, for a much greater range (about tenfold), and the same array configuration, this was not the case, indicating that the algorithm's parameters were not fully optimized for this particular function, i.e., the Bartlett processor for the

given simulation. This is a common pitfall when using any search algorithm such as genetic algorithms, simulated annealing, neural networks, hill climbers, and many others: When pitted against a black box environment, they have equivalent average behavior.[25] Research on adaptive algorithms incorporating as much knowledge as is available is most likely to be fruitful.

One method to overcome the problem based on *a priori* information on the environment is to place the hydrophone array at the appropriate depth so that the most important lower-order modes are detected at the array. This approach has been used in test case 2, where the turning depth of the propagating modes has been taken into account.

On the other hand, in the range-dependent case, the time for the experimental tuning of the algorithms parameters is very high, even within the adiabatic approximation and another approach had to be studied. In the framework of the present work, another genetic algorithm was developed, appropriate for maintaining enough diversity in the population, so that it could contain all the peaks of the fitness landscape, with minimum effort in the adjustment of its parameters. Furthermore, it is an easily parallelizable algorithm, using as many processors working in parallel, as half the population size![26]

II. TREATMENT OF THE RANGE DEPENDENCE

The recovery of range-dependent structures in the sea environment has always been a subject of research in ocean acoustic tomography. So far, good results of inversions have been reported with all ocean tomography methods with both simulated and real data for realistic range-independent environments. The same is true for cases when an average sound-speed profile is to be recovered. The problem of the recovery of range-dependent features is more difficult and theoretical justification of the existence of a unique solution to this inverse problem at a vertical slice is still lacking. However, it has been shown that a range-dependent reference environment in connection with a modal inversion scheme can lead to the reconstruction of a range-dependent inhomogeneity such as a sea eddy if the structure has a known compact support.[6,12] In this case, the recovery of the range-dependent structure has been achieved only because a range-dependent reference environment has been determined and there was an indication of the location of the eddy. The scheme cannot work with a range-independent reference environment.

Attempts to apply matched-field processing in range-dependent environments have been reported with encouraging results when a tomography experiment in which several pairs of sources and receivers are considered has been simulated.[8,27]

In the work presented here, MFP has been applied for a simulated experiment incorporating a single pair of sources–receivers, defining a vertical slice. The search space is relatively wide and does not take into account range-dependent features. However, physical constraints of the structure to be recovered are considered. Bearing in mind that in this case several mathematically acceptable solutions might appear but the procedure deals with the recovery of a sea eddy, its smoothness has been used as the critical feature and no sound-speed fluctuations following temperature fluctuations in the horizontal are allowed. This condition leads to the discarding of all nonphysically acceptable solutions.

III. THE MODAL-PHASE APPROACH

Linear inversion approaches have been used in ocean acoustic tomography when good *a priori* knowledge of the environment under study is ensured. In these cases, a ''reference'' environment can be defined with parameters exhibiting little difference with respect to them of the ''actual'' one. Their differences are further associated with differences of measurable quantities of the acoustic field due to a known source in the actual environment, with respect to calculations of the same quantities for the reference environment.

This concept was initially applied in ocean acoustic tomography with ray theory, the measured quantities being the arrival times of the acoustic energy packs propagating along the acoustic rays.[1] Later, Zaitsev[4] and Shang[2,5] used the normal-mode representation of the acoustic pressure and proposed the modal phase or the modal travel times to be the measured quantities, depending on whether an array of hydrophones or a single hydrophone is available as measuring device. Recently, Taroudakis showed that in the adiabatic approximation, both approaches can be used for recovery of range-dependent structures of compact support such as eddies at a vertical slice.[6,12] The recoverable quantity is the difference of the sound-speed profile, with respect to the reference one:

$$\delta c(r',z)=c(r',z)-c_0(r',z). \tag{4}$$

The question of how close to the actual environment the reference one should be so that the results obtained by the linear scheme are acceptable is still unresolvable. Although it is well understood that the parameters of the two environments should not differ significantly, so that the linear approach is valid, it has been shown that in some cases the method can work well, even when there is a relatively great difference among them.[6]

When empirical orthogonal functions (EOFs) can be used to represent the sound-speed variation, in the form

$$\delta c(r',z)=\sum_{\lambda=1}^{\Lambda} P^{\lambda}(z)q^{\lambda}(r'), \tag{5}$$

it can be shown[6] that the perturbation formula associating the coefficients $q^{\lambda}(r')$ with the measured modal-phase difference is

$$\delta\phi_n(r)=-\int_0^r \sum_{\lambda=1}^{\Lambda} \frac{1}{k_0^n(r')} \int_0^{\infty} \rho_0^{-1}(r') \times P^{\lambda}(z)|u_n^0(z;r')|^2 k_0^2(r',z)\,\frac{q^{\lambda}(r')}{c_0(r',z)}\,dz\,dr'. \tag{6}$$

Here, $\delta\phi_n(r)$ is the modal-phase difference calculated from the measurements of the modal phase after suitable mode filtering,[28]

$k_0(r',z)$ is the wave number at range r' and depth z,

ρ_0 is the density on water for the reference environment,

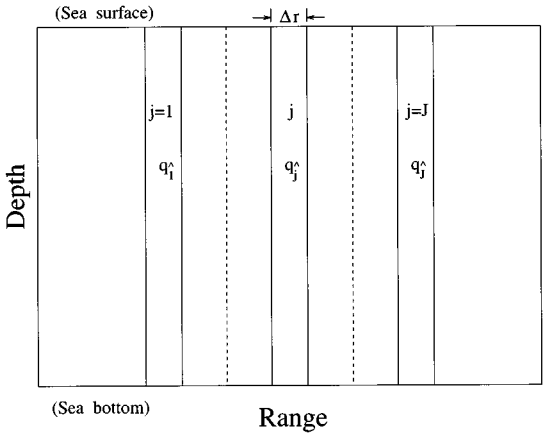


FIG. 1. Discretization of the field.

$u_n^0(z;r')$ is the eigenfunction of order n for the ''depth'' problem in range r',

$k_n^0(r')$ is the associated eigenvalue,

and we have assumed that the sound-speed variation can be well represented using Λ orders of EOFs.

It is recalled, that modal phase is defined as the integral:

$$\phi_n(r) = \int_0^r k_n(r')dr' = \overline{k_n} r. \tag{7}$$

Discretizing the region supporting the eddy in J vertical segments of width Δr (see Fig. 1), in each one of which the coefficients $q^{\lambda}(r')$ are constant $(q_j^{\lambda}, j=1,2,\ldots,J)$, we approximate the integration with the summation:

$$\delta\phi_n(r) = \sum_{j=1}^{J} \sum_{\lambda=1}^{\Lambda} Q_{na,j}^{0,\lambda}\, q_j^{\lambda} \Delta r, \tag{8}$$

where $Q_{na,j}^{0,\lambda}$ is a kernel matrix calculated for the reference environment.

A linear system is thus defined:

$$\delta\phi = \mathbf{Q_a^0} \cdot \mathbf{q}, \tag{9}$$

where $\delta\phi$ is a vector with N elements calculated using the acoustic measurements, $\mathbf{Q_a^0}$ is a known $N \times M$ matrix ($M = J \cdot \Lambda$) calculated for the background environment, and $\mathbf{q}$ is the vector containing the M unknown coefficients of the EOFs. The system can be solved using any suitable method. Note, however, that the condition of the matrix $\mathbf{Q_a^0}$ is not always good and special treatment should be applied for a reliable solution of the overdetermined system (9).[6,8] In our case, we use a singular value decomposition method to invert for the vector $\mathbf{q}$ and obtain from it the actual sound-speed structure.

Note that a similar approach based on perturbation formula (9), has been proposed by Rajan *et al.* for the identification of bottom properties[29] and that in our approach we have assumed that EOFs themselves do not change with range. However, it would be interesting to consider cases where range-dependent EOFs are defined.

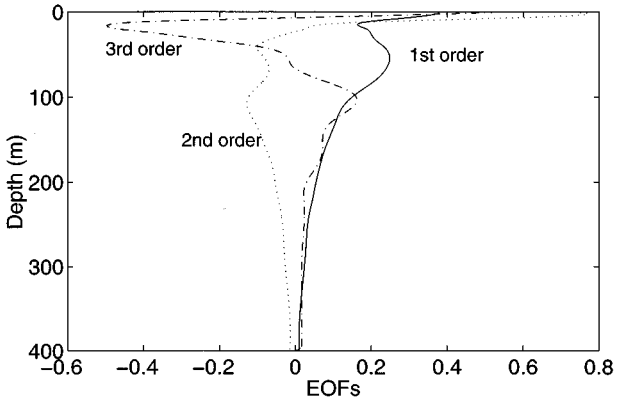


FIG. 2. The empirical orthogonal functions (EOFs) of the test case I.

IV. THE HYBRID APPROACH: MODAL-PHASE TOMOGRAPHY IN CONNECTION WITH MATCHED-FIELD PROCESSING

Since the environment obtained by matched-field processing is very close to the actual one (but not identical to it), it can serve as the reference environment for the application of a modal inversion scheme, which in turn will ''fine tune'' the results obtained through matched-field tomography.

Accordingly, a hybrid approach is developed in two complimentary stages. The first stage consists of the application of a matched-field algorithm with a wide search space to provide a good approximation of the environment to be recovered. The second stage consists of the application of a linear modal approach with the reference environment that was calculated by the matched-field algorithm.

In order that the hybrid scheme is applicable in realistic situations, it is necessary that both inversion algorithms should be based on data taken by the same measuring device. Since the matched-field processing utilizes a vertical array of hydrophones, the modal-phase approach is used instead of modal travel time, since modal-phase data can be obtained as a side product of the output of the vertical array following a filtering procedure.

It is interesting to note that the application of the hybrid scheme is justified only in cases that the results obtained by matched-field processing cannot be considered as satisfactory. As will be shown in the next section, among the three cases studied, only the range-dependent environment was eligible for the application of this scheme. In the range-independent cases results obtained by matched-field processing are acceptable by themselves and no improvement is expected by fine tuning.

V. THE TEST CASES

A. Test case I: Range-independent environment—Shallow water

The first test case appeared in Refs. 21 and 30 and consists of a shallow water environment of 400 m depth, with a semi-infinite bottom of density 1.5 g/cm^3 and compressional speed of 1600 m/s. The sound-speed profile in this environment is described in terms of three orders of empirical orthogonal functions as in Fig. 2. The coefficients of the EOFs

TABLE I. EOF coefficients and search space for the environment of test case I.

Order	EOF coefficient	Bounds	
		Lower	Upper
1	−44.713	−81.92	0
2	44.438	0	81.92
3	−14.892	−20.48	20.48

appear in Table I together with their upper and lower bounds, determining their search space.

A vertical array of hydrophones consisting of 20 sensors spanning the first 75 m in depth, is considered as the measuring device and the processing is made at the frequency of 200 Hz. The source has been positioned at the depth of 50 m. The environment is taken to be noisy.

Since the sensors in the array are spaced a half-wavelength apart, noise may be modeled to first order as being spatially white. Matching fields were computed using the normal-mode code SNAP.

The selection process chosen here for the genetic algorithm is remainder stochastic sampling without replacement, which minimizes stochastic errors in sampling. Selection pressure was introduced by linear fitness scaling, to prevent premature convergence to suboptimal solutions, as well as a meaningless wandering among mediocre solutions in the end.[21]

Typically, population size, crossover, and mutation rates are interrelated. Their choice might be a difficult nonlinear problem by itself. We preferred to keep a small population size (50) to speed up the convergence rate, along with a 100% crossover rate, as recommended for the abovementioned selection procedure.[22] Two-point crossover was performed in the case of low signal-to-noise ratio (SNR$=-5$ dB), while for the higher SNR of 10 dB we preferred multiple-point crossover. This was justified empirically but it can be explained as well: the lower the SNR, the noisier the Bartlett processor, and the larger the diversity in the population; there is no reason to increase it by performing multiple-point crossover.

Since mutation is useful to keep a continuing ''genetic''

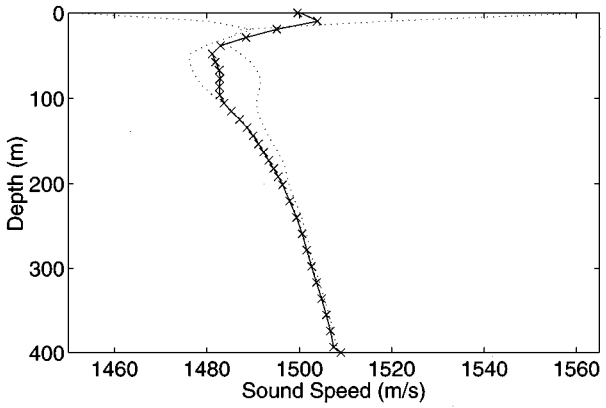


FIG. 3. Actual and recovered sound-speed profiles for the test case I. The solid line represents the actual profile, the line with *x* the recovered profiles, while dotted lines present the limits of the search space.

TABLE II. Recovered EOF coefficients using MFP for test case I.

Order	SNR$=10$		SNR$=-5$	
	Mean value	st. dev.	Mean value	st. dev.
1	−43.645	3.027	−44.555	0.966
2	43.648	3.780	44.546	1.331
3	−15.106	1.425	−14.865	1.352

diversity but should also be controlled, we introduced its rate among parameters for iteration. The algorithm therefore adaptively modifies the mutation operator as the evolution proceeds, and the mutation probability varies between 0.1% and 4%. This would be interesting in the case of nonstationary object functions, where the fitness landscape changes over time.

In small populations, stochastic errors in sampling result in ultimate convergence on one ''peak'' or another without differential advantage, sometimes missing the true peak. To reduce the effects of these errors and enable stable subpopulations to form around each peak, we induce niche and ''species'' formation. A power-law sharing function, based on the relative bit difference, is used to define the neighborhood and degree of sharing for each string in the population. Similar strings derate one another's fitness values and limit the uncontrolled growth of particular species within a population. Speciation has been further exploited through methods of mating restriction. When ''mating'' occurs at random, high performance but dissimilar parents tend to generate low-performance ''off-springs'' (lethals).[22] To cause *similar* individuals (i.e., belonging to the same niche) to be recombined, various approaches have been examined; the algorithm used here chooses them according to their Euclidean distance, i.e., it selects the ''nearest neighbors,'' as long as the average performance increases; otherwise, random recombination is performed temporarily.

The convergence rate of the algorithm has been studied by the authors for two different SNRs, 10 and -5 dB, respectively, and the GA has shown to be relatively insensitive to the decrease of the SNR.[21] Actually, in the case of white noise, the true peak always corresponds to the maximum value of Bartlett processor; increased noise levels simply produce large numbers of local optima that, instead of divert-

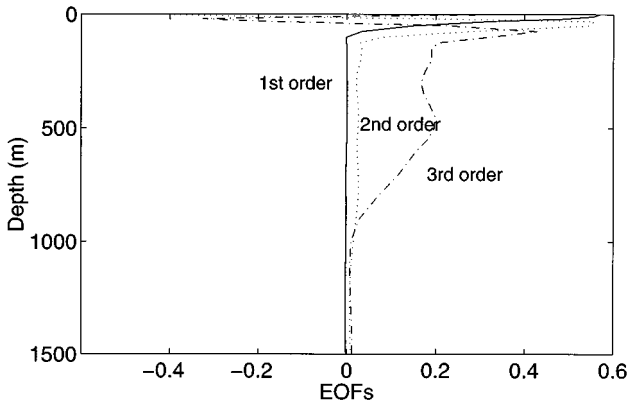


FIG. 4. The EOFs of the test case II.

TABLE III. EOF coefficients and search space for the environment of test case II.

Order	EOF coefficient	Bounds	
		Lower	Upper
1	18.48	−40.96	40.95
2	2.88	−40.96	40.94
3	0.96	−20.48	20.46

ing the algorithm from further improvement, are exploited to improve performance, while the search for more global optima goes on. All runs have been performed on a SUN SPARC10 workstation. Most of the CPU time is consumed in solving the forward problem for each model's evaluation; thus it is proportional to the total number of models evaluated, which can vary between 1000 and 4000, due to the different random seed used in every run of the algorithm.

Figure 3 shows actual and recovered SSPs, as well as the upper and lower bounds of the search space, for the matched-field inversion, where the recovered profile corresponds to the best fit of the models. Statistical analysis of the best models obtained from 20 trials for the cases of SNR = 10 and − 5 dB gives the mean values and variances appearing in Table II, which should be compared with the true values of Table I. The agreement is remarkably good.

B. Test case II: Range-independent environment—Deep water

The second test case corresponds to a real environment taken from the THETIS-II experiment, which was a feasibility study of the application of ocean acoustic tomography for monitoring the Western Mediterranean basin.[31] Range-averaged sound-speed profiles had been recovered using single receivers at specified sites in the area.[3] We chose the longer path in the area (550 km) and performed a study of the applicability of MFP for the recovery of this profile. MFP was performed at 250 Hz which was the central frequency of the low-frequency source utilized in the experiments, and the single receiver was replaced in our simulations by an array of hydrophones in two different configurations described below. The source has been considered in its actual depth of 150 m, and an absorbing bottom was considered.

TABLE IV. Recovered EOF coefficients using MFP for test case II.

Order	415 hydrophones	20 hydrophones
1	22.06	15.20
2	0.46	3.54
3	2.98	0.54

The sound-speed profile difference with respect to a linear one (1503 m/s in the sea surface, 1547 m/s at the sea bottom) is represented in terms of the three orders of empirical orthogonal functions, which appear in Fig. 4.

We have chosen the average profile determined by CTD casts from March 1992 as the structure to be recovered. The coefficients of these EOFs appear in Table III, together with upper and lower bounds used in the MFP to define the search space. The average profile appears in Fig. 5.

The first array simulated in the present study consisted of 415 elements spanning half the water depth. The results of the MFP are presented in Table IV and Fig. 5 and are clearly considered satisfactory.

Then, trying to minimize the number of sensors and determine the minimum useful aperture, we studied the modal content of the pressure field in the environment, based on information of the sound-speed profile measured at the receivers location. The study was realized using ray-mode analogy, for the association of grazing angles with each normal mode at a given depth. Additionally, the modal energy distribution for each receiver was calculated and evanescent modes at the various depths were defined. The results showed that the greater part of the acoustic energy, corresponding to the lower-order modes (10), propagate at depths spanning from 50 to 300 m. At the same time, EOFs analysis in this area confirmed that the sound-speed variation described by the two first modes is important up to a depth of 250 m. Combination of these two observations leads to the conclusion that an adequate estimation of the sound-speed structure by acoustic inversion could be obtained by placing the hydrophone array to span the major part of this depth.

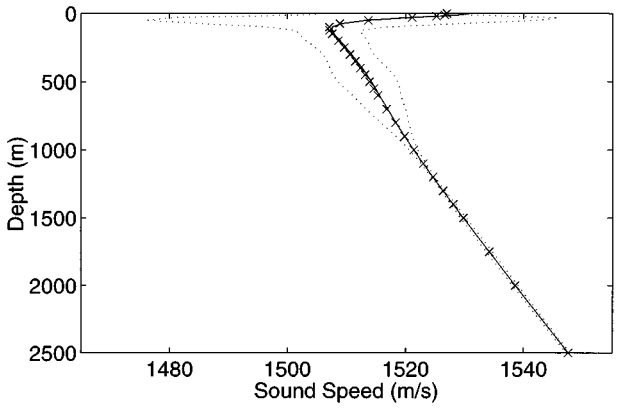


FIG. 5. Actual and recovered sound-speed profiles for the test case II. Simulations with the vertical array of 415 hydrophones. The solid line represents the actual profile, the line with *x* the recovered profiles, while dotted lines present the limits of the search space.

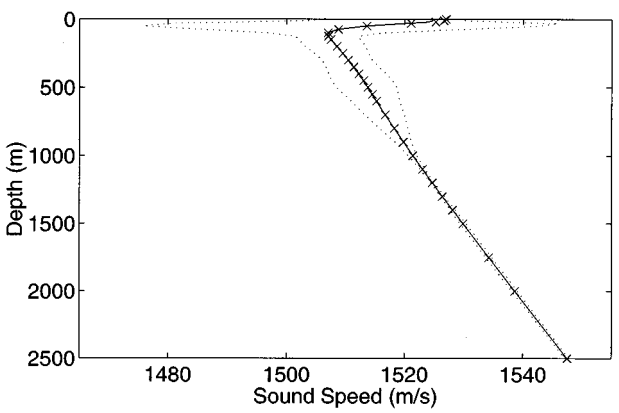


FIG. 6. Actual and recovered sound-speed profiles for the test case II. Simulations with the vertical array of 20 hydrophones. The solid line represents the actual profile, the line with *x* the recovered profiles, while dotted lines present the limits of the search space.

TABLE V. The sound-speed profile in the environment of the test case III, outside the eddy area.

Depth (m)	Sound speed (m/s)
0	1500
100	1495
400	1509

To test the hypothesis we simulated an array of just 20 hydrophones placed 10 m apart to each other and spanning the depth from 70 down to 260 ms below the sea surface.

The genetic algorithm used here was the same as in the previous case. The application of the MFP algorithm using the SNAP code to calculate the complex acoustic field, resulted in the coefficients of Table IV obtained with white noise of 10-dB SNR. The reconstructed profile shown in Figs. 5 and 6 exactly fits the actual one, proving once more the good performance of the MFP for the estimation of sound-speed profiles in range-independent environments, and the validity of the hypothesis made regarding the reliability of the reduced-length array as the measuring device for MFP.

C. Test case III: An environment with a cold eddy

The third environment is a range-dependent one. The water depth and bottom properties are exactly the same as in the test case I. The background sound-speed structure is determined by the simple sound-speed profile of Table V. We introduce a smooth cold eddy (Fig. 7) at a vertical slice confined between 2000 and 3200 m. The figure has been reproduced by describing the eddy by five sound-speed profiles in equidistant segments spanning the 2000–3200 m range. The sound-speed profiles in the eddy have been expanded again in terms of the EOFs of the first environment and the corresponding coefficients appear in Table VI. Figure 8 presents the sound-speed profiles of the five segments.

We consider a vertical array of 28 sensors spanning the first 105 m in depth, and a single frequency source (200 Hz) in an environment characterized by white noise with SNR=10 dB.

Table VI presents upper and lower bounds, and the precision of the coding of each EOF coefficient, which are kept the same for all segments.

In problems presenting multiple solutions, like the one at hand, one has to use a search algorithm capable of finding most of them. Physical constraints, or some other criteria, can be considered afterward to exclude the unacceptable solutions, as we do here.

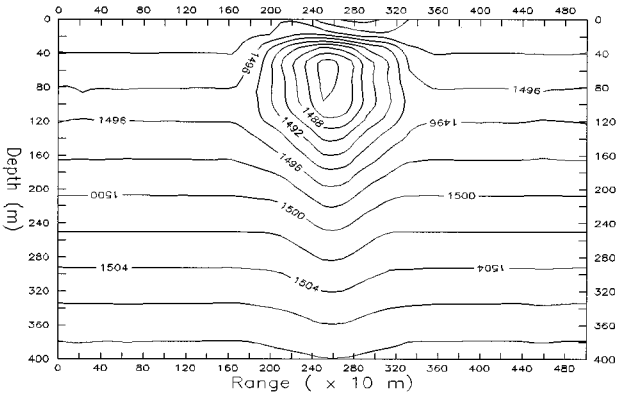


FIG. 7. The environment of the test case III.

A traditional GA will, at best, ultimately converge to a population containing only one of these solution. Niching mechanisms permit the forming and maintaining of stable subpopulations, or niches, within a single population, corresponding to different peaks of the fitness landscape, and, thus, different solutions.[22] Sharing, used in cases I and II, and crowding have proven effective niching methods. Since the given problem has many extraneous peaks that are similar in fitness to the desired peaks, sharing without a proper fitness scaling would allocate too many individuals to them.[32] In order to save time, we avoided the experimental adjustment of the fitness scaling, along with the other GA parameters—population size (n), and discretization step—and we preferred deterministic crowding (DC),[26] described in the Appendix, which is simple, fast, and parallel.

In the first stage of inversions dealing with MFP only, both synthetic data and replica fields were computed using the adiabatic version of the normal-mode model KRAKEN.[33]

The results obtained by the MFP algorithm appear in Figs. 9 and 10, the former presenting contours of equal sound speed, the latter presenting sound-speed profiles of the five segments, together with the profiles of the ''actual'' environment. To obtain the results presented in Table VII, which correspond to the recovered structure of Fig. 9, the average of the best 20 points in the final population (size $n = 200$), has been derived having already discarded the solutions that did not conform to a single eddy environment, according to the restrictions described in Sec. II.

The solution obtained by MFP is satisfactory. The shape of the eddy is well represented. However, the representation of the sound-speed field close to the surface is not accurate. This behavior can be attributed to the relatively large deviations of the recovered coefficients of the second EOFs with respect to the actual ones (cf. Tables VI–VII).

Since the deviation from the actual environment was not

TABLE VI. EOF coefficients and search space for the environment of test case III.

	Segment					Bounds		
Order	1st	2nd	3rd	4th	5th	Lower	Upper	Precision
1	−19.21	−33.007	−44.713	−25.657	− 8.726	−76.2	0.0	0.6
2	27.845	34.352	44.438	32.824	22.876	0.0	75.6	1.2
3	−11.105	−11.008	−14.892	−13.073	−12.001	−37.2	0.0	1.2

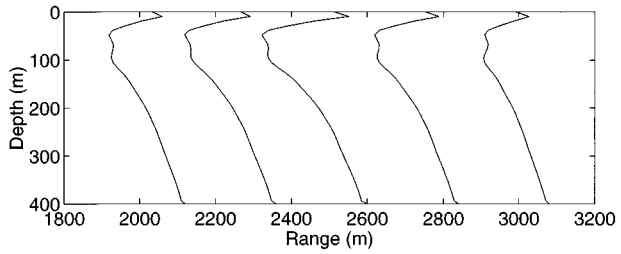


FIG. 8. The sound-speed profiles of the five segments that describe the eddy of the test case III.

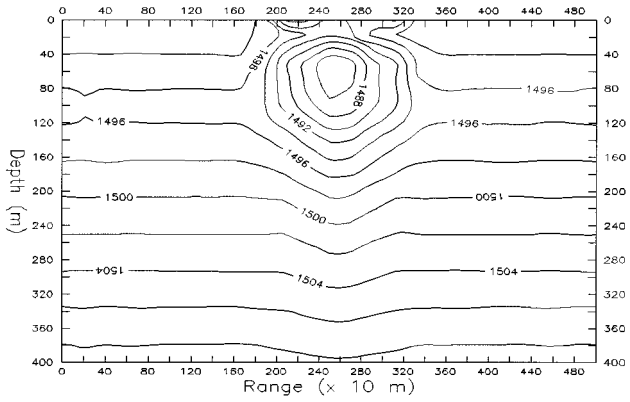


FIG. 9. The environment of test case III, reconstructed by the MFP.

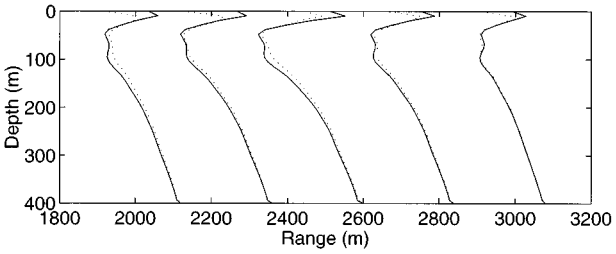


FIG. 10. The sound-speed profiles describing the eddy of the test case III. Comparison of the actual (solid lines) and recovered profiles (dotted lines) using matched-field inversion.

TABLE VII. Recovered EOF coefficients using MFP for test case III.

	Segment				
Order	1st	2nd	3rd	4th	5th
1	−20.715	−36.765	−45.345	−26.325	−13.575
2	11.760	18.180	31.200	22.200	14.580
3	−10.980	−13.860	−10.080	−14.340	−13.080

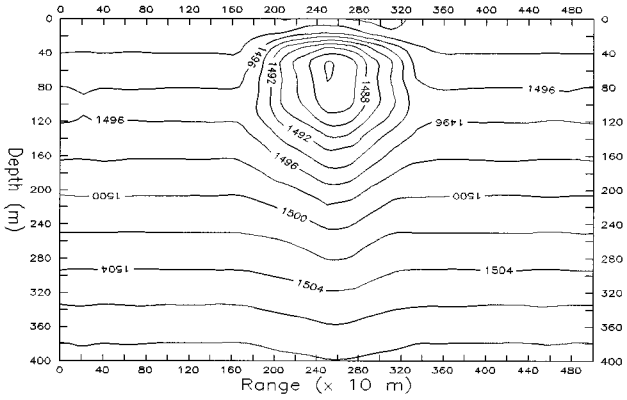


FIG. 11. The environment of test case III, reconstructed by means of the hybrid scheme. Background environment determined by the MFP. Fine tuning using the modal-phase approach.

considerable, the modal-phase approach has been used at a second stage to obtain a better approximation of the actual eddy. As already stated in Sec. IV, results of the MFP inversions served as the background environment for the modal-phase algorithm. We considered the modal phase of the first 10 modes to be known, obtained by suitable mode filtering.

The application of the inversion algorithm using again the KRAKEN code to calculate the modal-phase differences with respect to the reference environment, resulted in sound-speed profiles for the five segments defining the structure appearing in Fig. 11. It is obvious that the eddy is now much better represented. The individual sound-speed profiles appear in Fig. 12 and the recovered EOF coefficients in Table VIII.

From the practical point of view, it is interesting to compare the final solution of this hybrid procedure with same obtained by modal phase approach only, as in Ref. 6. The eddy is equally well reproduced in both cases, however in Ref. 6, good *a priori* estimate of the EOF coefficients has been considered (a rather optimistic case). Here, no reference environment has been used. Instead we considered a very wide search space using MFP. The whole procedure is now realistic and may well represent an actual situation in a tomographic experiment with a vertical array of hydrophones serving as the measuring device. It remains to study the reliability of the modal-phase calculations which is a task that has already been undertaken by the authors.

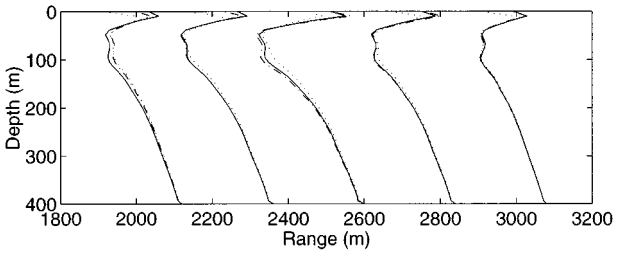


FIG. 12. The sound-speed profiles describing the eddy of the test case III. Comparison of the actual (solid) background (dotted) and recovered profiles (dashed-dotted) using the hybrid scheme.

TABLE VIII. Recovered EOF coefficients using the hybrid approach for test case III.

	Segment				
Order	1st	2nd	3rd	4th	5th
1	−16.3283	−35.0108	−46.8045	−24.2481	−10.0163
2	21.0388	31.5324	47.6640	34.4505	23.3496
3	−10.3013	−11.9814	−17.6835	−14.8792	−11.4820

VI. CONCLUSIONS

The use of a genetic algorithm in a matched-field processing scheme for the reconstruction of range-independent shallow and deep water environments, as well as a shallow water environment with a cold eddy, has been shown to be successful. In all these cases the sound-speed structure is described in terms of sound-speed EOFs. For the range-dependent case especially, although the problem is ill-posed and multiple solutions could be admissible within the limits of the MFP, the choice of the averaging over them resulted in satisfactory reconstruction of all the environments studied, starting from wide search spaces. For this case, the physical constraint of the smoothness of its structure has also been imposed as an additional restriction factor.

The results of the MFP for the case of the sea eddy have been proven to provide a reliable reference environment for the application of a modal-phase inversion scheme for a substantial improvement of the inversions, under the assumption of the existence of reliable measurements of the modal phase for a number of propagating modes. This hybrid inversion procedure seems to be a promising inversion tool for ocean acoustic tomography when measurements are performed with the help of a vertical array of hydrophones. It remains to assess the reliability of the modal-phase measurements which is also an important factor for the overall evaluation of the scheme.

From the computational point of view, the structure of the algorithm is such that parallel processing is allowed at various stages. Thus an appreciable reduction of execution time is achieved, and the whole procedure can be completed at reasonable time.

APPENDIX

The procedure followed for the application of deterministic crowding (DC) with threshold accepting (TA) is briefly described:

All population elements mate randomly each generation, undergoing single-point crossover with probability $p_c = 1$, and mutation with probability $p_m = 0.001$. The ''children'' compete with nearest ''parents'' (we use their Euclidean distance as measure d) which are replaced if they are inferior. This is called replacement selection and it is the only selection that occurs in DC. Following a suggestion in Ref. 28, we use a mechanism to control convergence. Instead of simulated annealing controls proposed in Ref. 26, we use a threshold accepting method (TA)[34] that has a simpler structure. TA includes ''children'' in the new population even if they are worse than their ''parents,'' but not much; the threshold that determines the maximum acceptable difference between new and old solutions, lowers progressively every generation until it becomes zero.

The algorithm is synoptically presented below:

(A) Set threshold $T > 0$ to a sufficiently high value.

(B) Initialize the population (randomly).

(C) Repeatedly generate each new population from the current population as follows:

(1) Do $n/2$ times:

(a) Select two ''parents,'' p_1 and p_2, randomly, without replacement

(b) Cross them, yielding ''children'' c_1 and c_2

(c) Apply mutation/other operators, yielding c_1' and c_2'

(d) If $[d(p_1,c_1') + d(p_2,c_2')] \leqslant [d(p_1,c_2') + d(p_2,c_2')]$:

- If $f(c_1') > f(p_1) - T$ replace p_1 with c_1';
- If $f(c_2') > f(p_2) - T$ replace p_2 with c_2';

Else

- If $f(c_2') > f(p_1) - T$ replace p_1 with c_2';
- If $f(c_1') > f(p_2) - T$ replace p_2 with c_1'.

(2) Periodically lower T till zero.

Since all population elements can proceed simultaneously, parallelism is limited only by population size n. We could use up to $n/2$ parallel processors via population decomposition.

[1] W. Munk and C. Wunsch, ''Ocean acoustic tomography: A scheme for large scale monitoring,'' Deep Sea Res. **26a**, 123–161 (1979).

[2] E. C. Shang and Y. Y. Wang, ''On the possibility of monitoring El Nino by using modal ocean acoustic tomography,'' J. Acoust. Soc. Am. **91**, 136–140 (1992).

[3] E. K. Skarsoulis, G. A. Athanassoulis, and U. Send, ''Ocean acoustic tomography based on peak arrivals,'' J. Acoust. Soc. Am. **100**, 797–813 (1996).

[4] V. Yu. Zaitsev, A. G. Nechaev, and L. A. Osrovskii, ''An algorithm for three-dimensional mode tomography of the ocean,'' Sov. Phys. Acoust. **33**, 657–658 (1987).

[5] E. C. Shang, ''Ocean acoustic tomography based on adiabatic mode theory,'' J. Acoust. Soc. Am. **85**, 1531–1537 (1989).

[6] M. I. Taroudakis and J. S. Papadakis, ''A modal inversion scheme for ocean acoustic tomography,'' J. Comput. Acoust. **1**, 395–421 (1993).

[7] A. Tolstoy, O. Diachok, and L. N. Frazer, ''Acoustic tomography via matched field processing,'' J. Acoust. Soc. Am. **89**, 1119–1127 (1991).

[8] A. Tolstoy, ''Linearization of the matched field processing approach to acoustic tomography,'' J. Acoust. Soc. Am. **91**, 781–787 (1992).

[9] A. B. Baggeroer, W. A. Kuperman, and P. N. Mikhalevsky, ''An overview of matched field methods in ocean acoustics,'' IEEE J. Ocean Eng. **18**, 401–424 (1993).

[10] M. B. Porter and A. Tolstoy, ''The matched field processing benchmark problems,'' J. Comput. Acoust. **2**, 161–185 (1994).

[11] S. 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).

[12] M. I. Taroudakis, ''A comparison of the modal-phase and modal-travel time approaches for ocean acoustic tomography'' in *Proceedings of the 2nd European Conference on Underwater Acoustics*, edited by L. Bjørnø (EC, Luxemburg, 1994), pp. 1057–1062.

[13] A. Tolstoy, *Matched Field Processing for Underwater Acoustics* (World Scientific, Singapore, 1993).

[14] A. Tolstoy, ''Using matched field processing to determine shallow water environmental parameters,'' in *Theoretical and Computational Acoustics*, edited by D. Lee, Y-H. Pao, M. H. Schultz, and Y.-C. Teng (World Scientific, Singapore, 1996), pp. 283–295.

[15] W. Menke, *Geophysical Data Analysis: Discrete Inverse Theory* (Academic, New York, 1989).

[16] T. C. Yang, ''A method of range and depth estimation by modal decomposition,'' J. Acoust. Soc. Am. **82**, 1736–1745 (1987).

[17] T. C. Yang, ''Effectiveness of mode filtering: A comparison of matched-field and matched mode processing,'' J. Acoust. Soc. Am. **87**, 2072–2084 (1990).
[18] P. Gerstoft, ''Inversion of seismoacoustic data using genetic algorithms and a posteriori probability distributions,'' J. Acoust. Soc. Am. **95**, 770–782 (1994).
[19] P. Gerstoft, ''Global inversion by genetic algorithms for both source position and environmental parameters,'' J. Comput. Acoust. **3**, 251–266 (1994).
[20] Z.-H. Michalopoulou, H. Martynov, and M. B. Porter, ''Simulated annealing and genetic algorithms for broadband source focalization,'' in *Proceedings of the 3rd European Conference on Underwater Acoustics*, edited by J. S. Papadakis (FORTH-IACM, Heraklion, 1996), pp. 409–414.
[21] M. I. Taroudakis and M. G. Markaki, ''Matched field ocean acoustic tomography using genetic algorithms,'' in *Acoustical Imaging*, edited by P. Tortoli and L. Masotti (Plenum Press, New York, 1996), pp. 601–606.
[22] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning* (Addison-Wesley, Reading, MA, 1989).
[23] F. B. Jensen and M. C. Ferla, ''SNAP: The SACLANTCEN normal-mode acoustic Propagation Model,'' SACLANTCEN Memorandum SM-121, La Spezia, Italy, 1979.
[24] E. Vavalis and C. Georgalakis, ''Normal-mode acoustic propagation models on distributed/parallel computer systems,'' FORTH/IACM Technical Report No. 94-9, 1994.
[25] H. Kargupta and D. E. Goldberg, ''SEARCH: An alternate perspective toward blackbox optimization,'' Department of Computer Science and Illinois Genetic Algorithms Laboratory, University of Illinois at Urbana-Champaign, 1995.
[26] S. W. Mahfoud, *A Comparison of Parallel and Sequential Niching Methods* (LBS Capital Management, Inc., Clearwater, FL, 1995).
[27] A. Tolstoy, ''Simulated performance of acoustic tomography via matched field processing,'' J. Comput. Acoust. **2**, 1–10 (1994).
[28] E. L. Lo, J. X. Zhou, and E. C. Shang, ''Normal-mode filtering in shallow water,'' J. Acoust. Soc. Am. **74**, 1833–1836 (1983).
[29] S. Rajan, J. F. Lynch, and G. V. Frisk, ''Perturbative inversion methods for obtaining bottom geoacoustic parameters in shallow water,'' J. Acoust. Soc. Am. **82**, 998–1017 (1987).
[30] G. A. Athanassoulis, J. S. Papadakis, E. K. Skarsoulis, and M. I. Taroudakis, ''A comparative study of two wave-theoretic inversion schemes in ocean acoustic tomography,'' in *Full Field Inversion Methods in Ocean and Seismo-Acoustics*, edited by O. Diachok *et al.* (Kluwer Academic, Dordrecht, 1995), pp. 127–132.
[31] U. Send U., G. Krahmann, D. Mauuary, Y. Desaubies, F. Gaiilard, T. Terre, J. Papadakis, M. Taroudakis, E. Skarsoulis, and C. Millot, ''Acoustic observations of heat content across the Mediterranean Sea,'' Nature (London) **385**, 615–617 (1997).
[32] S. W. Mahfoud, ''Crowding and preselection revisited,'' in *Parallel Problem Solving from Nature*, edited by R. Manner and B. Manderick (Elsevier, New York, 1992), pp. 27–36.
[33] M. B. Porter, ''The KRAKEN normal mode program,'' SACLANT Memorandum SM-245, 1991.
[34] G. Dueck and T. Scheurer, ''Threshold accepting: A general purpose optimization algorithm appearing superior to simulated annealing,'' J. Comput. Phys. **90**, 161–175 (1990).

High-resolution analysis of the complex wave spectrum in a cylindrical shell containing a viscoelastic medium. Part I. Theory and numerical results

Johannes Vollmann[a] and Jürg Dual
Institute of Mechanics, ETH Zürich, Swiss Federal Institute of Technology, CH-8092 Zürich, Switzerland

(Received 19 December 1996; accepted for publication 7 April 1997)

In this study the relation between frequency and complex wave number of axisymmetric wave modes in an isotropic, thin-walled, cylindrical shell containing a linear viscoelastic medium is derived. Shell wall bending and longitudinal motion are coupled in an empty cylindrical shell. When a viscoelastic medium is enclosed, the shell motion is affected by the complex bulk and shear modulus, as well as by the density of the medium enclosed. A Maxwell model is used for both complex Lamé constants λ and μ to describe the constitutive equations of the medium. By varying the complex moduli, the medium can be modeled as an inviscid fluid, an elastic material, or anything between these two extremes. The interaction of the thin-walled linear elastic shell and the viscoelastic medium is discussed numerically by calculating the complex dispersion relation. Numerical results are presented for an empty shell and a shell filled with three types of core material: an inviscid fluid, a shear dissipative fluid, and a shear elastic fluid. In a companion paper [J. Vollmann *et al.*, J. Acoust. Soc. Am. **102**, 909–920 (1997)], the experimental setup and the signal processing used to perform the high-resolution measurement of the dispersion relation are described in detail. Theoretical and experimental results are compared. [S0001-4966(97)01108-9]

PACS numbers: 43.35.Mr, 43.40.Ey, 43.40.Cw, 43.20.Tb [HEB]

INTRODUCTION

The motivation of this research project is the physical interpretation of the interaction of an elastic shell and a viscoelastic core with the aim of determining viscoelastic material properties by measuring wave propagation phenomena in a broad frequency range. Enclosed in a tube, the viscoelastic medium is uncomplicated to handle and if the tube itself is acting as a sensor, this technique can easily be integrated into an on-line chemical process. The various aspects of problems relating to wave propagation in a tube filled or surrounded by a fluid have been the subject of numerous studies. Most of these studies pertain to the water hammer effect or to the propagation of heartbeat in human arteries. Despite the large number of papers devoted to this subject (see below), the phenomenological interpretation as well as the experimental verification, especially of the higher-order modes in the complex space (complex wave number versus frequency), are rather limited.

The history of analytical description of waves in cylindrical structures dates back to 1876, when Pochhammer[1] published his work. Early work on the water hammer effect was performed by Joukowsky[2] and in 1955, Skalak[3] published an extension of the water hammer theory. A brief overview of further significant contributions to this subject is given by Fuller and Fahy.[4] The problem of a thin-walled elastic shell with a compressible, inviscid fluid was first solved by Lin and Morgan[5] in which the results were limited to real wave numbers. Assuming the fluid to be inviscid, Fuller and Fahy[4] calculated the dispersion curves for axisymmetric waves ($n=0$) and for the first-order nonaxisymmetric modes ($n=1$) in the plane of real wave number versus frequency, the plane of imaginary wave number versus frequency, as well as in the space of complex wave number versus frequency. Guo presented approximate solutions to the dispersion equation for fluid-loaded shells[6] and has studied the attenuation of helical waves in a three-layered shell consisting of elastic and viscoelastic layers.[7] On the experimental side, Plona *et al.*[8] analyzed waves in a thick-walled tube which was surrounded and/or filled with water and compared it with the theoretical results from Sinha *et al.*[9] Recently, Lafleur and Shields[10] discussed the influence of tube material on the first two modes in a liquid-filled tube, considering long wavelength only.

In a field ''somewhere'' between acoustics, fluid mechanics, and solid mechanics, many effects and phenomena interfere with each other, and the borders between these classical disciplines fade with the increasing generality of the constitutive equations. The various motivations, the complexity of the interaction of two waveguides, and the lack of powerful computers led to diverse assumptions and simplifications in the past. In this study, a general viscoelastic approach for dilatation and shear deformation of the core enables one to calculate the dispersion relation of an empty shell, a fluid-filled shell, a shell filled with a viscoelastic medium, and a linear elastic rod with circular cross section by varying the values of the viscoelastic material properties only.

[a]Present address: Northwestern University, Center for Quality Engineering and Failure Prevention, 2137 N. Sheridan Road, Evanston, IL 60206-3020, Electronic mail: vollmann@nwu.edu

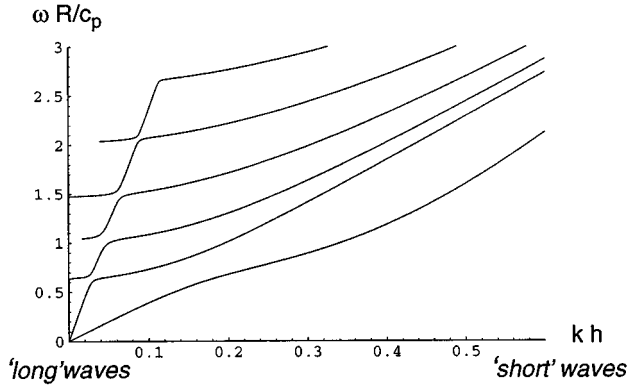


FIG. 1. ''Normalized frequency versus normalized wave number'' representational form of axisymmetric wave modes in a fluid-filled cylindrical shell.

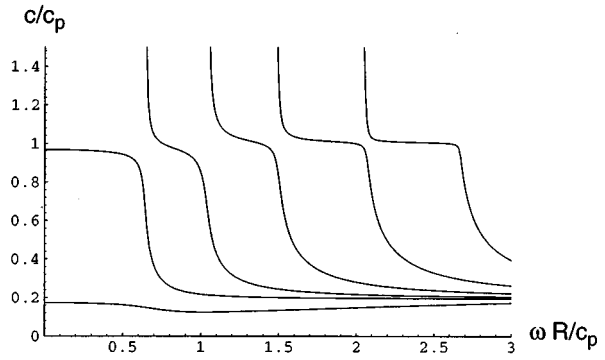


FIG. 2. ''Normalized phase velocity versus normalized frequency'' representational form of axisymmetric wave modes in a fluid-filled cylindrical shell.

I. NATURE AND REPRESENTATIONAL FORMS OF DISPERSION

Interference between P, SH, and SV waves and their reflections in a three-dimensional medium of finite cross-sectional dimensions causes so-called ''wave modes'' of guided waves in structures. In a limited interval of frequency and wave number, a finite number of propagating modes regularly exist. In the case of axisymmetric waves in a cylindrical structure, it is sufficient to consider one-dimensional wave propagation in the direction of the symmetry axis (x axis) in order to characterize a three-dimensional structure. The propagating harmonic waves can be described as

$$u(x,t)=\sum_{n=1}^{n=p}\left(A_n e^{i(\omega t+k_n x)}+B_n e^{i(\omega t-k_n x)}\right), \tag{1}$$

in which u denotes the displacement of a surface point of the structure and p denotes the number of traveling modes in the x direction. Here ω indicates the frequency and k signifies the complex wave number which can be regarded as ''frequency'' in space domain. The imaginary part of the wave number k indicates a decaying or a growing wave mode depending on its sign, whereas the complex constants A_n and B_n describe the amplitude and initial phase.

Dispersion diagrams referring to a steel shell with a diameter of 26 mm, a wall thickness of 0.5 mm, and filled with silicone oil are presented here to illustrate different representational forms of dispersion. Detailed information about the material properties and physical interpretations of the shapes of the curves are given in Sec. III.

Regarding Eq. (1), the clearest way to represent a dispersion relation is by drawing curves in the ωk plane, as shown in Fig. 1. Here R denotes the mean radius of the shell, h the wall thickness, and c_p the velocity of longitudinal waves in a plate [see Eq. (A1) in the Appendix]. However, other forms like phase velocity versus frequency (Fig. 2) or group velocity versus frequency (Fig. 3) are also very common. All diagrams shown in Figs. 1–3 represent the same dispersion relation, thus providing clarity for those readers who are not familiar with the ωk form.

Since the experimental verification directly yields ω vs k, the theoretical dispersion curves will also be represented in the frequency–wave-number plane or in the space defined by real frequency and complex wave number in the forthcoming sections.

II. THEORETICAL MODEL

A. Dynamic shell equations

A model for isotropic thin-walled elastic shells, including effects of shear deformation and rotary inertia, forms the basis of the dispersion relation. Despite the fact that torsional modes are axisymmetric as well, they are not considered in this investigation. Thus the dynamic equilibrium of an infinite shell element bordering the viscoelastic core is formulated in the plane defined by the radial and the longitudinal coordinate. In Fig. 4 the forces acting on a differential shell element are shown.

$$N_{x,x}=\rho_s h u_{x,tt}+\sigma^f_{rx}, \tag{2}$$

$$Q_{r,x}-\frac{N_\vartheta}{R}=\rho_s h u_{r,tt}+\sigma^f_{rr}, \tag{3}$$

$$-M_{x,x}+Q_r=\rho_s I\xi_{,tt}+\sigma^f_{rx}\frac{h}{2}. \tag{4}$$

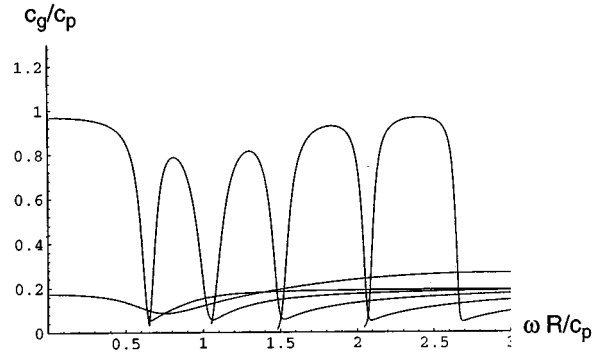


FIG. 3. ''Normalized group velocity versus normalized frequency'' representational form of axisymmetric wave modes in a fluid-filled cylindrical shell.

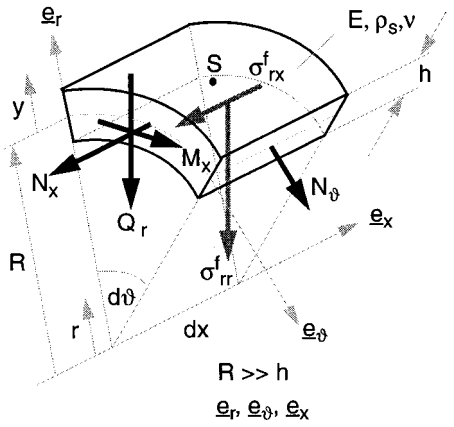


FIG. 4. Differential shell element and forces involved in axisymmetric wave propagation.

In Eqs. (2)–(4) σ^f denotes the stress at the interface between the inner-shell surface and the viscoelastic core medium. I indicates the moment of inertia, and ξ the slope of the cross section due to bending. In general, index s designates terms associated with the shell and index f terms associated with the core. Derivatives are written in the short form: $u_{r,x}$ denotes the derivative of the radial displacement u_r with respect to x.

According to Timoshenko,[11] effects of shear deformation and rotary inertia of the linear elastic shell are considered. Here the ''wrong'' assumption that the cross section remains plane is corrected by the factor κ. Values for κ depend on the shape of the cross section and are chosen from a publication by Cowper.[12] The limits of the authors' approach were evaluated by comparing the results for an empty shell with those of an exact, three-dimensional theory presented by Herrmann and Mirsky.[13] The difference between the exact and the approximate theory amounted to less than 0.2% for the frequency range considered.

Figure 5 and Eq. (5) show that the derivative of the deflection u_r with respect to the longitudinal coordinate x consists of a part caused by bending (ξ) and a part caused by shear deformation (γ_0):

$$u_{r,x}=\xi+\gamma_0. \tag{5}$$

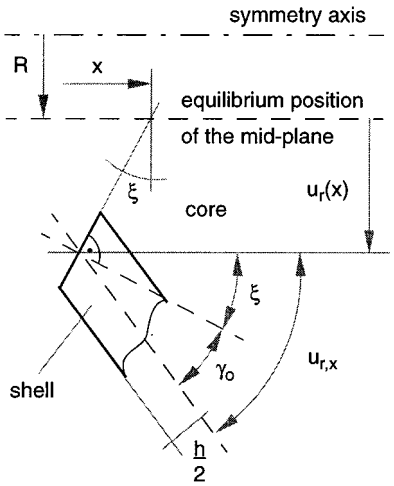


FIG. 5. Deflection of the cross section of a shell element due to bending and shear deformation.

Using the kinematic relations for cylindrical coordinates, Eqs. (6) and (7), and the constitutive equations for linear elastic shell material, Eqs. (8) and (9), the stress resultants M, N, and Q are obtained by integrating the stresses through the thickness of the shell element [see Eqs. (10)–(13)]:

$$\epsilon_x=u_{x,x}-\xi y, \tag{6}$$

$$\epsilon_\vartheta=\frac{u_r}{R}, \tag{7}$$

$$\sigma_x=\frac{E}{1-\nu^2}\,(\epsilon_x+\nu\epsilon_\vartheta), \tag{8}$$

$$\sigma_\vartheta=\frac{E}{1-\nu^2}\,(\epsilon_\vartheta+\nu\epsilon_x). \tag{9}$$

Here, y indicates the integration variable in the shell wall in radial direction, and higher-order terms are neglected:

$$N_x=\int_{h/2}^{h/2}\sigma_x\,dy=\frac{Eh}{1-\nu^2}\left(u_{x,x}+\nu\,\frac{u_r}{R}\right), \tag{10}$$

$$N_\vartheta=\int_{h/2}^{h/2}\sigma_\vartheta\,dy=\frac{Eh}{1-\nu^2}\left(\frac{u_r}{R}+\nu u_{x,x}\right), \tag{11}$$

$$Q_r=hG\kappa(u_{r,x}-\xi) \tag{12}$$

$$M_x=\int_{h/2}^{h/2}y\sigma_x\,dy=-\frac{Eh^3}{12(1-\nu^2)}\,\xi_{,x}. \tag{13}$$

Substituting the expressions for the stress resultants M, N, and Q into Eqs. (2)–(4) and eliminating ξ, leads to Eqs. (14) and (15), a system of two coupled differential equations. In these equations D denotes the bending stiffness of a plate $[D=Eh^3/(12(1-\nu^2))]$ and C denotes the longitudinal stiffness of a plate $[C=Eh/(1-\nu^2)]$. It can easily be seen that in the limit $R\to\infty$, Eqs. (14) and (15) are decoupled and represent the dynamic equations for bending waves and longitudinal waves in a thin plate, respectively,

$$\begin{aligned}
&Du^s_{r,xxxx}-\left(\frac{D\rho_s}{\kappa G}+\rho_s I\right)u^s_{r,ttxx}\\
&\quad-\frac{D}{\kappa Gh}\left[\sigma^f_{rr,xx}+\frac{C}{R}\left(\frac{u^s_{r,xx}}{R}+\nu u^s_{x,xxx}\right)\right]\\
&\quad+\left(\rho_s h+\frac{C\rho_s I}{R^2\kappa Gh}\right)u^s_{r,tt}+\frac{C}{R}\left(\frac{u^s_r}{R}+\nu u^s_{x,x}\right)\\
&\quad=-\sigma^f_{rr}+\sigma^f_{rx,x}\,\frac{h}{2}-\frac{\rho_s I}{\kappa Gh}\left[\sigma^f_{rr,tt}+\rho_s h u^s_{r,tttt}+\frac{C}{R}\,\nu u^s_{x,ttx}\right],
\end{aligned} \tag{14}$$

$$C\left(u^s_{x,xx}+\nu\,\frac{u^s_{r,x}}{R}\right)=\rho_s h u^s_{x,tt}+\sigma^f_{rx}. \tag{15}$$

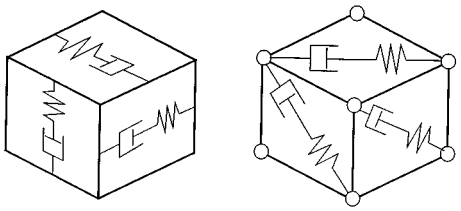


FIG. 6. Illustration of Maxwell behavior for dilatation and shear deformation.

The next step in deriving the theory is the formulation of the interface stress σ_{rr}^{f} and σ_{rx}^{f} as a function of shell displacement.

B. Wave propagation in the linear viscoelastic cylindrical core

To describe the interface stresses between the shell and the viscoelastic core, wave propagation in the three-dimensional core must be investigated. For harmonic time functions, the constitutive equation for isotropic linear viscoelastic material is derived from the equations for linear elastic material, written in terms of Lamé constants λ and μ as follows:

$$\sigma_{ij}=\lambda\,\delta_{ij}\epsilon_{kk}+2\mu\epsilon_{ij}. \tag{16}$$

By adding loss moduli to the storage moduli, the whole investigation becomes complex:

$$\sigma_{ij}=\lambda^{*}(\omega)\,\delta_{ij}\epsilon_{kk}+2\mu^{*}(\omega)\epsilon_{ij}. \tag{17}$$

In Eq. (17), λ^{*} and μ^{*} are complex, frequency-dependent functions describing the relation between stress σ_{ij} and strain ϵ_{ij}. The Kronecker symbol is denoted by δ_{ij}. In general, the asterisk designates values which are complex due to viscoelasticity.

In this investigation, Maxwell models are used for both $\lambda^{*}(\omega)$ and $\mu^{*}(\omega)$, represented in Eqs. (18) and (19), in order to describe the viscoelastic behavior of the medium. As illustrated in Fig. 6, this model consists of spring and damper on the same axis subjected to the same force. For high frequency, compared with the reciprocal value of the relaxation time T, this model is governed by the spring and for low frequency by the damper:

$$\lambda^{*}(\omega)=\frac{\lambda_0 i\omega}{i\omega+1/T_{\lambda}}, \tag{18}$$

$$\mu^{*}(\omega)=\frac{\mu_0 i\omega}{i\omega+1/T_{\mu}}. \tag{19}$$

Equation (19) shows the complex shear modulus μ^{*} and Eq. (20) the complex bulk modulus B^{*}:

$$B^{*}(\omega)=\lambda^{*}(\omega)+\tfrac{2}{3}\mu^{*}(\omega). \tag{20}$$

Equation (21) represents the complex shear viscosity and Eq. (22) the complex bulk viscosity:

$$\eta^{*}(\omega)=\frac{\mu^{*}(\omega)}{i\omega}, \tag{21}$$

$$b^{*}(\omega)=\frac{\lambda^{*}(\omega)}{i\omega}+\frac{2}{3}\frac{\mu^{*}(\omega)}{i\omega}. \tag{22}$$

A Maxwell model is one of the simplest models used to characterize viscoelastic material behavior. However, if experimental results show reason to use a more sophisticated model, this can easily be realized by modifying Eqs. (18) and (19).

The displacement field is expressed in terms of derivatives of potentials, as shown in Eq. (23). This form decomposes the displacement field in a dilatational part, associated with φ, and a rotational part, associated with $\boldsymbol{\Psi}$:

$$\mathbf{u}^{f}=\boldsymbol{\nabla}\varphi+\boldsymbol{\nabla}\times\boldsymbol{\Psi}. \tag{23}$$

Since the motion is assumed to be axisymmetric and torsional modes are not considered in this investigation, all derivatives with respect to ϑ vanish and the vector field $\boldsymbol{\Psi}$ is reduced to one dimension, as shown in the following equation:

$$\boldsymbol{\Psi}=\begin{pmatrix}0\\ \Psi\\ 0\end{pmatrix}. \tag{24}$$

The approach used in Eq. (23) enables one to separate the dynamic equation of motion into two so-called Helmholtz equations:

$$\rho_f\varphi_{,tt}-(\lambda^{*}(\omega)+2\mu^{*}(\omega))\nabla^2\varphi=0; \tag{25}$$

$$\rho_f\Psi_{,tt}-\mu^{*}(\omega)\nabla^2\Psi=0. \tag{26}$$

The potentials φ and Ψ propagate with the velocities c_{φ}^{*} and c_{Ψ}^{*}, as specified in Eqs. (27) and (28) where k_{φ}^{*} and k_{Ψ}^{*} are the corresponding wave numbers. In linear elasticity, the common designations for c_{φ}^{*} and c_{Ψ}^{*} and c_1 and c_2:

$$c_{\varphi}^{*}=\sqrt{\frac{\lambda^{*}(\omega)+2\mu^{*}(\omega)}{\rho_f}},\quad k_{\varphi}^{*}=\frac{\omega}{c_{\varphi}^{*}}, \tag{27}$$

$$c_{\Psi}^{*}=\sqrt{\frac{\mu^{*}(\omega)}{\rho_f}},\quad k_{\Psi}^{*}=\frac{\omega}{c_{\Psi}^{*}}. \tag{28}$$

To solve the Helmholtz equations (25) and (26), the following assumptions are made for waves propagating along the tube:

$$\varphi(r,x,t)=f(r)e^{i(\omega t-k^{*}x)}, \tag{29}$$

$$\Psi(r,x,t)=g(r)e^{i(\omega t-k^{*}x)}. \tag{30}$$

The derivatives of the potential functions, Eqs. (29) and (30), are expressed in terms of Laplace operators for cylindrical coordinates, Eqs. (31) and (32). Substituting these expressions into the Helmholtz equations, the differential equations (33) and (34) for the unknown functions $f(r)$ and $g(r)$ are obtained:

$$\nabla^2\varphi=\varphi_{,xx}+\frac{1}{r}\,\varphi_{,r}+\varphi_{,rr}, \tag{31}$$

$$\nabla^2\boldsymbol{\Psi}=\begin{pmatrix}0\\ \Psi_{,rr}+\frac{1}{r}\Psi_{,r}+\Psi_{,xx}-\frac{1}{r^2}\Psi\\ 0\end{pmatrix}, \tag{32}$$

$$r^2 f_{,rr}+r f_{,r}+r^2(k_\varphi^{*^2}-k^{*2})f=0, \tag{33}$$

$$r^2 g_{,rr}+r g_{,r}+[r^2(k_\Psi^{*^2}-k^{*2})-1]g=0. \tag{34}$$

For simplicity of notation, the following abbreviations are used:
$\beta_\varphi^{*^2}=k_\varphi^{*^2}-k^{*2}$ and $\beta_\Psi^{*^2}=k_\Psi^{*^2}-k^{*2}$.
According to Abramowitz and Stegun,[14] Eqs. (33) and (34) are solved by combinations of complex Bessel functions J_0, J_1, K_0, K_1, represented in Eqs. (35) and (36), in which a_1, a_2, b_1, b_2 are arbitrary constants to be determined by satisfying boundary and interface conditions:

$$f(r)=a_1 J_0(\beta_\varphi^* r)+a_2 K_0(i\beta_\varphi^* r), \tag{35}$$

$$g(r)=b_1 J_1(\beta_\Psi^* r)+b_2 K_1(i\beta_\Psi^* r), \tag{36}$$

The displacement field can now be expressed by substituting Eqs. (29) and (30) into the following equation:

$$\begin{aligned} &u_r=\varphi_{,r}-\Psi_{,x}, \quad u_\vartheta=0,\\ &u_x=\varphi_{,x}+\Psi_{,r}+\frac{1}{r}\Psi. \end{aligned} \tag{37}$$

If the motion is axisymmetric, the radial motion vanishes on the symmetry axis and the axial motion must remain finite. Since the values of the Bessel functions K_0 and K_1 go to infinity for r going to zero, this condition can only be satisfied if the values of the constants a_2 and b_2 are set to zero.

The kinematic relations for the axisymmetric case [Eq. (38)],

$$\begin{aligned} &\epsilon_{rr}=u_{r,r}, \quad \epsilon_{\vartheta\vartheta}=u_r/r,\\ &\epsilon_{xx}=u_{x,x}, \quad \epsilon_{rx}=\tfrac{1}{2}(u_{r,x}+u_{x,r}) \end{aligned} \tag{38}$$

combined with the stress–strain relations [Eq. (17)], represent the final step in formulating the problem.

C. Boundary and interface conditions between shell and core

At the interface between the shell and the viscoelastic core ($r=R-h/2$), the radial displacements in each medium are equal and the components of the stress vector are equal with opposite directions. The stress functions in the cylindrical core, $\sigma_{rr}^f(r)$ and $\sigma_{rx}^f(r)$, are presented in Eqs. (39) and (40):

$$\begin{aligned} \sigma_{rr}^f(r)=e^{i(\omega t-k^* x)}\Big\{&-J_0(\beta_\varphi^* r)[a_1 k^{*2}\lambda^*+a_1\beta_\varphi^{*^2}(\lambda^*+2\mu^*)]\\ &+J_1(\beta_\varphi^* r)\left[\frac{2a_1\mu^*\beta_\varphi^*}{r}\right]+J_0(\beta_\Psi^* r)[2b_1 ik^*\beta_\Psi^*\mu^*]\\ &-J_1(\beta_\Psi^* r)\left[\frac{2b_1 ik^*\mu^*}{r}\right]\Big\}, \end{aligned} \tag{39}$$

$$\begin{aligned} \sigma_{rx}^f(r)=e^{i(\omega t-k^* x)}\mu^*\{&J_1(\beta_\varphi^* r)2a_1 ik^*\beta_\varphi^*+J_1(\beta_\Psi^* r)\\ &\times[b_1(k^{*2}-\beta_\Psi^{*^2})]\}. \end{aligned} \tag{40}$$

For the shell displacement, the following time- and space-harmonic approaches are made [Eqs. (41) and (42)]:

$$u_r^s=a_r e^{i(\omega t-k^* x)}=u_r^f(r=R-h/2), \tag{41}$$

$$u_x^s=a_x e^{i(\omega t-k^* x)}. \tag{42}$$

Whereas the radial displacement of the shell midplane equals the radial displacement at the core interface, the longitudinal displacement at the interface consists of both the longitudinal shell displacement and the displacement caused by the rotation of the cross section due to shell wall bending [see Eq. (43) and Fig. 5]:

$$u_x^f(r=R-h/2)=u_x^s+\xi\frac{h}{2}=a_x e^{i(\omega t-k^* x)}+\xi\frac{h}{2}. \tag{43}$$

The deflection component ξ of the cross section caused by bending is a function of wave speed of lateral waves and is expressed in terms of a_r by Eqs. (5)–(13).

D. Dispersion relation

Equations (41) and (43) show the interface condition between shell and core. Substituting these equations into Eq. (37) provides two linear equations for the arbitrary constants a_x, a_r, a_1, and b_1. Two further equations are obtained by substituting Eqs. (39) and (40) into the dynamic shell equations (14) and (15):

$$\begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14}\\ m_{21} & m_{22} & m_{23} & m_{24}\\ m_{31} & m_{32} & m_{33} & m_{34}\\ m_{41} & m_{42} & m_{43} & m_{44}\end{bmatrix}\begin{pmatrix} a_x\\ a_r\\ a_1\\ b_1\end{pmatrix}=\begin{pmatrix}0\\0\\0\\0\end{pmatrix}. \tag{44}$$

This homogeneous system of linear equations for the constants a_x, a_r, a_1, and b_1 [Eq. (44)] has a nontrivial solution if the determinant of the coefficients equals zero. (The coefficients m_{ij} are presented in the Appendix 6.)

$$\mathrm{Det}[m_{ij}]=0. \tag{45}$$

Since this condition contains ω, k^*, as well as all geometrical and material parameters of the shell and the viscoelastic core, Eq. (45) represents the implicit dispersion relation. Assuming a real ω, the corresponding wave numbers k_n^* are calculated with a complex root-finding algorithm based on Newton's method. Then ω is increased by $\Delta\omega$ and the wave numbers of the former step are used as initial values for the next root-finding procedure.

E. Normalized shell displacement

A knowledge of the shell displacement as a function of frequency is of importance in interpreting the shapes of the

various modes. As the shell displacement can be measured in only one direction, this knowledge is also very helpful in determining the optimal angle between the shell axis and the measuring laser beam.

In general, the relative shell displacement, i.e., the relation between u_x and u_r, is complex. Hence it contains information of the amplitude and the relative phase shift between u_x and u_r of continuously propagating wave modes. In Eq. (46) the complex relative displacement relation is written in terms of the coefficients m_{ij} found in the Appendix:

$$\frac{a_x}{a_r}=\frac{(m_{42}m_{13}-m_{43})(m_{24}m_{13}-m_{23}m_{14})+(m_{23}-m_{22}m_{13})(m_{44}m_{13}-m_{43}m_{14})}{-m_{13}m_{41}(m_{24}m_{13}-m_{23}m_{14})+m_{13}(m_{44}m_{13}-m_{43}m_{14})}. \tag{46}$$

Since a_x and a_r are complex constants, the shell displacements [Eqs. (41) and (42)] can be written as follows [Eqs. (47) and (48)]:

$$u_r^s=|a_r|e^{i(\omega t-k^*x)+i\alpha_r}, \tag{47}$$

$$u_x^s=|a_x|e^{i(\omega t-k^*x)+i\alpha_x}. \tag{48}$$

Now the complex displacement relation can be separated into the amplitude ratio, Eq. (49), and the phase angle between u_x and u_r, Eq. (50):

$$\frac{|u_x^s|}{|u_r^s|}=\frac{|a_x|}{|a_r|}=\left|\frac{a_x}{a_r}\right|, \tag{49}$$

$$\Delta\alpha=(\alpha_x-\alpha_r)=\arg\left(\frac{a_x}{a_r}\right). \tag{50}$$

Numerical analysis of Eq. (50) shows that the phase angle is very sensitive to small differences in the viscoelastic material properties.

F. Normalized stress field in the viscoelastic core

Another important characteristic feature of the wave modes is their associated stress field. To discuss the penetration depth of the wave propagation in radial direction, the normalized stress vectors $\tilde{\sigma}_{rr}^f$ and $\tilde{\sigma}_{rx}^f$ versus a normalized radial coordinate $\tilde{r}$ are calculated for one particular mode at one particular frequency:

$$\tilde{r}=\frac{r}{R-h/2}. \tag{51}$$

Transforming the variable r to $\tilde{r}$ in Eqs. (39) and (40) leads to Eqs. (52) and (53). Again, these expressions are complex. The real part represents a snapshot of the stress field in the core, the absolute value represents the maximum reached at a specific location during one cycle, and the phase angle between the stress in the core and the stress at the shell wall is associated with the time delay of the stress when propagating in radial direction through the viscoelastic material:

$$\tilde{\sigma}_{rr}^f(\tilde{r})=\frac{\sigma_{rr}^f(\tilde{r})}{\sigma_{rr}^f(\tilde{r}=1)}, \tag{52}$$

$$\tilde{\sigma}_{rx}^f(\tilde{r})=\frac{\sigma_{rx}^f(\tilde{r})}{\sigma_{rx}^f(\tilde{r}=1)}. \tag{53}$$

Equations (52) and (53) enable one to calculate the normalized partial stress due to one traveling mode and are tools for the interpretation of the mode shapes. However, the real stress situation in a system excited with one particular frequency consists of the superposition of the stress fields caused by all traveling modes at that frequency.

III. SOME EXAMPLES OF NUMERICAL RESULTS

In this research project, the theoretical formulation and the experimental investigation began at the same time and were carried out simultaneously. Therefore the experimental feasibility has a strong influence on the selection of the parameters being varied in this section. Most of the experiments are conducted using a steel tube ($E=2.000\times10^{11}$ N/m^2, $\nu=0.28$, $\rho_s=7800$ kg/m^3) with a diameter of 26 mm and a wall thickness of 0.5 mm, so $R/h=26$. This parameter, as well as the shell material, will remain unchanged in the investigation below. One of the main interests is the influence of the shear behavior of the core on the complex dispersion relation. The following aspects may help to clarify the influence of the various material parameters on the shape of the dispersion curves.

In general, wave propagation in a medium demands elasticity to store potential energy and inertia to store kinetic energy. Therefore discussing wave propagation in a viscoelastic material deals necessarily with the principal limits of wave propagation. The core material can either be dominantly elastic or dominantly viscous for dilatation and/or shear deformation. A viscous behavior for dilatation suppresses the propagation of P waves which leads to dispersion curves similar to those of an empty shell. Since no such material was found in reality, this class will not be discussed.

First, the dynamic behavior of an empty shell is discussed as done by Sinha *et al.*[9] This approach may help one to understand the dispersion relation of a shell with a viscoelastic core as a blend of two interacting wave guides.

Second, the numerical discussion is divided into three main sections according to the nature of the shear resistance of the core:

(i) A shell filled with an inviscid core material (corresponding to an lossless fluid).
(ii) A shell filled with a shear dissipative core material (corresponding to a viscous fluid).
(iii) A shell filled with a shear elastic core material.

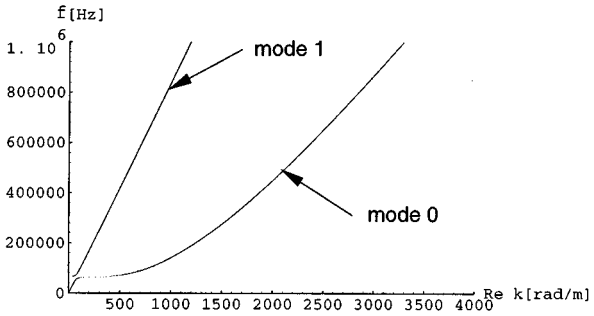


FIG. 7. Dispersion diagram of an empty thin-walled shell.

In all cases, the real part of the velocity of P waves $c_\varphi^* = 1005.66$ m/s [see Eq. (27)] and the density of the core $\rho_f = 969.00$ kg/m^3 are kept constant, the complex bulk modulus [Eq. (20)] being dominantly elastic.

The starting point for this numerical investigation are values which have achieved the best fit of a measured dispersion relation of silicone oil with high viscosity at low shear velocity. However, to provide a nondimensional form, the material constants are normalized with the Young's modulus E_s, the shear modulus G_s, and the density ρ_s of the shell material. The relaxation times are normalized with the upper frequency $f_{\max}$ of the considered interval which equals from 1 kHz to 1.0 MHz (see Table I).

Due to numerical reasons, the values can neither be set to zero nor to infinity, but rather to very high or very low values. For example, a relaxation time T_λ of 1 s already signifies an ''infinitely'' long time in a frequency range between 1 kHz and 1 MHz, causing primarily elastic behavior.

A. The empty shell

For the model considered, the wave spectrum of axisymmetric wave modes in the empty cylindrical shell is comprised of three modes only. As mentioned earlier, the torsional mode is excluded from this investigation. The shape of the other modes is either lateral (radial), longitudinal, or both. The lack of material damping and radiation loss leads to purely real dispersion curves, as shown in Fig. 7. This dispersion diagram will be discussed beginning with the lowest frequency. For low frequencies and long wavelength, the waves propagate with the velocity of longitudinal waves in a bar c_0. As soon as the wavelength reaches the circumferential dimension, an additional mode (mode 1) appears. In that particular frequency range, a strong coupling between radial and longitudinal motion occurs for both modes. Whereas mode 0 changes its nature from dominantly longitudinal to radial, mode 1 arises as a breathing mode transforming to an almost purely longitudinal mode (see Fig. 8). The coupling between radial and longitudinal motion leads to the interaction of the core material with the shell through its shear and bulk resistance.

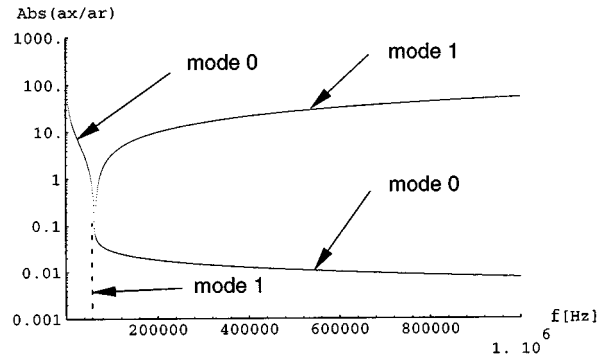


FIG. 8. Relation between axial and radial displacement of the empty shell.

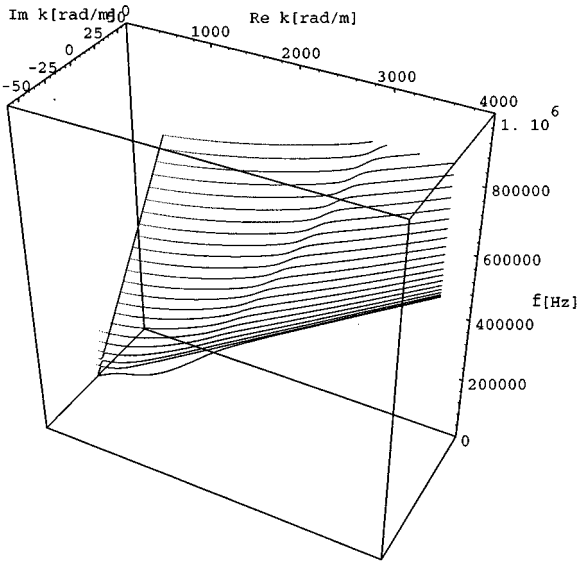


FIG. 9. Dispersion curves of the first 24 axisymmetric modes of a shell filled with an inviscid material (lossless fluid).

B. The shell filled with an inviscid core

In contrast to Fig. 7 in which the behavior of one single waveguide is presented, all of the following diagrams are determined by a dynamic interaction between shell and core.

Provided that no energy is emitted from the outer surface of the shell, the dispersion curves also remain in the plane of real wave number versus frequency when the shell is filled with an inviscid core material, as shown in Fig. 9. The dispersion curves are calculated from the right-hand side to the left, starting with the highest wave number. Following

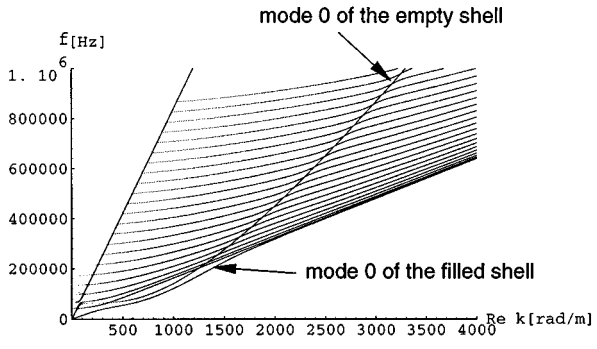


FIG. 10. Superposition of the dispersion curves for an empty shell and for a shell filled with an inviscid core material.

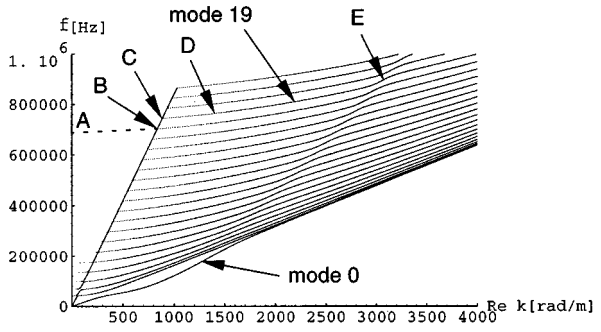


FIG. 11. Dispersion curves of the first 24 axisymmetric modes in the real plane.

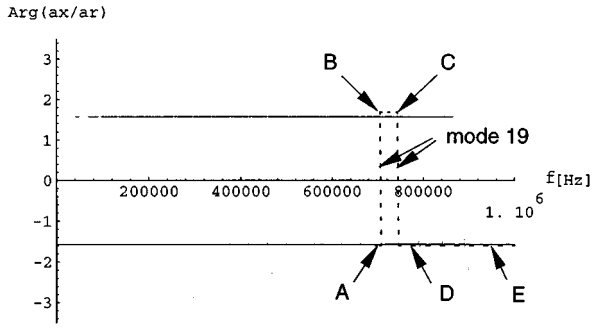


FIG. 13. Phase angle between a_x and a_r for the first 24 modes in a shell filled with inviscid material according to Eq. (50). The dashed line emphasizes mode 19.

the graph of one single mode, a small region of high slope appears. As the superposition of the diagrams for the empty shell and for the shell with inviscid core (Fig. 10) shows, this steeper region is linked with the flexural mode of the empty shell (mode 0 in Fig. 7). The straight line of the dispersion curves in Fig. 11 is associated with membrane waves in the shell propagating in longitudinal direction with the velocity c_p [see Eq. (A1) in the Appendix]. The correct shape of the curves on the left side of the straight line can be recognized when observing the first five modes. Except for mode 0 and mode 1, each mode has a cutoff frequency but due to the close proximity of the curves to each other, the root-finding algorithm jumps from one mode to the next lowest mode.

In order to explain the different stages of mode shapes along a dispersion curve of one graph, the letters A–E are added to the dispersion curve of mode 19 (Fig. 11), as well as to the corresponding graph in the displacement ratio diagram (Fig. 12).

Figure 12 shows the absolute value of the normalized complex shell displacement ratio in logarithmic scale versus frequency. Point A indicates the cutoff frequency of mode 19 where the shell is breathing simultaneously over its whole length and the motion is purely radial. The next cutoff frequency, in terms of growing frequency, occurs roughly as soon as the ratio formed by the core radius and a half P wavelength is an integer.

At point B the longitudinal component of the shell motion is equal to approximately 20 times the radial motion and remains almost constant in a small frequency range before it increases again, reaching its maximum in point C. Following the graph of mode 19 towards higher frequencies leads to point D where the radial motion is dominant again.

Since a_x and a_r are complex, the ratio a_x/a_r can be discussed in terms of amplitudes [see Fig. 12], as well as in terms of the relative time delay between the two components, expressed by the argument of the complex ratio. For a shell filled with an inviscid core, the phase between a_x and a_r jumps from $-\pi/2$ to $+\pi/2$ so quickly whenever a mode reaches the area of the membrane wave, that no vertical lines can be seen in Fig. 13. This situation changes drastically when the core material has shear resistance. To show the line of the phase graph of mode 19, a dashed line has been added to the diagram.

C. The shell filled with a dissipative core

The dissipative shear behavior of the core causes attenuation of the shell waves, which results in a negative value of the imaginary part of the wave number. In order to explain the topology of the dispersion curves of a shell filled with a dissipative core, the quantity δ, representing a dynamic penetration depth is introduced. Equation (54) shows that a wave motion in an arbitrary one-dimensional continuum, described by ψ and complex k^* with a negative imaginary part, can be

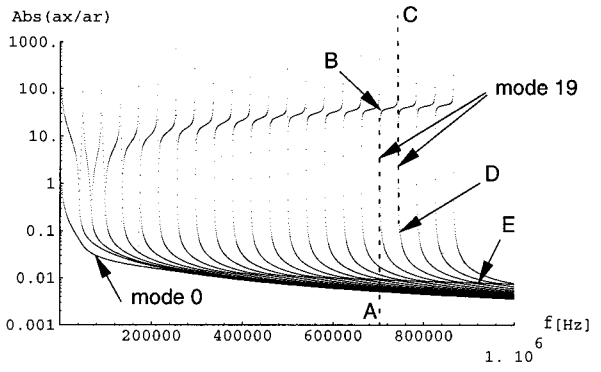


FIG. 12. Relation between axial and radial shell displacement of the first 24 axisymmetric wave modes in a cylindrical shell containing an inviscid material.

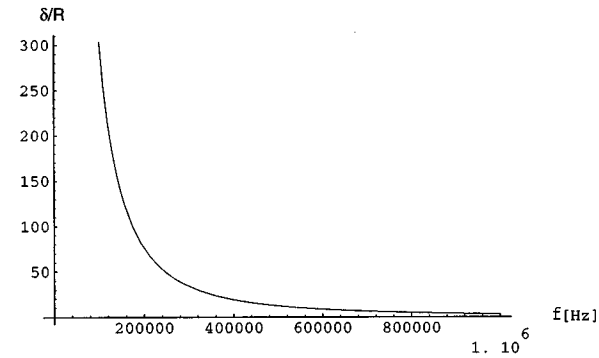


FIG. 14. Normalized 0.2% boundary layer characterizing the attenuation of P waves in the shear dissipative core material.

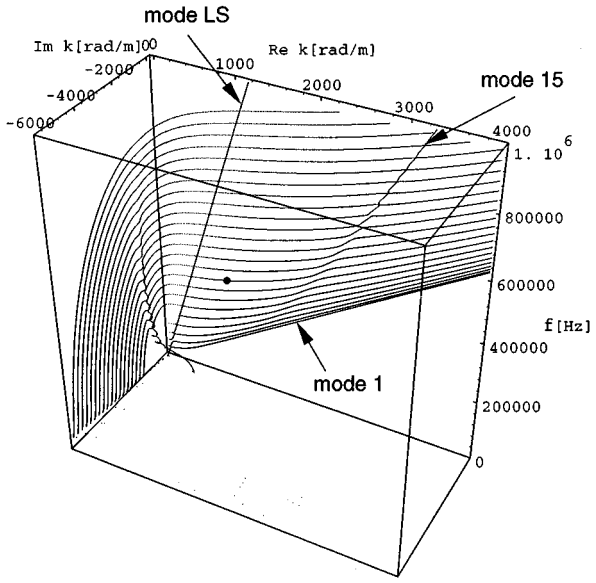


FIG. 15. Dispersion curves of axisymmetric modes of a shell filled with a shear dissipative core material (classical fluid).

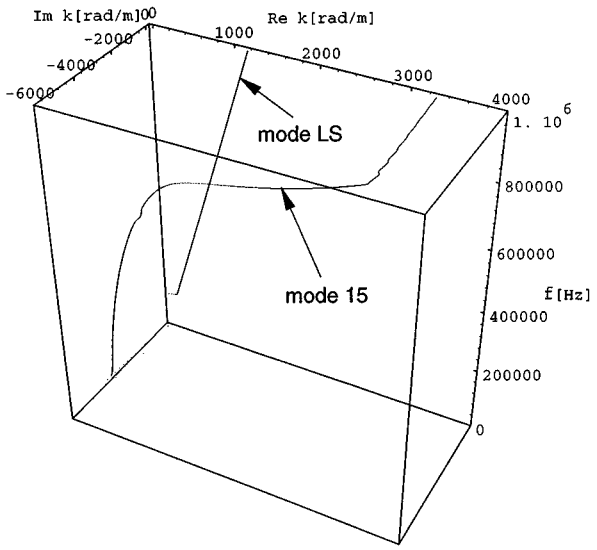


FIG. 17. Dispersion curves of the modes which show the decoupling of the shell and the core caused by the dissipative behavior of the core material.

written in terms of an undamped traveling wave times an exponential attenuation factor in the z direction.

$$u(z,t)=ae^{i(\omega t-k^*z)}=ae^{i(\omega t-k_r z)}e^{k_i z}, \tag{54}$$

$$\delta=\frac{2\pi}{|k_i|}. \tag{55}$$

According to Eq. (55), δ describes the distance in m from the emitting source to that point in the viscoelastic medium, where the amplitude has dropped to about 0.2% of the initial amplitude, thus providing an approximate value of the penetration depth of mechanical disturbances in an attenuating medium. Assuming z to be a radial coordinate in the core, outgoing from the inner-shell surface, the definition of δ is not valid any more due to the cylindrical shape of the core, for the following qualitative discussion, however, it still fits the purpose.

According to Eq. (20), the bulk modulus is a function of λ^* and μ^* and therefore a strong shear dissipative behavior primarily suppresses the propagation of S waves as well as the propagation of P waves. Figure 14 shows the normalized thickness of the boundary layer of P waves in the core material discussed in this section to be a function of frequency. One can see that above approximately 700 kHz, δ reaches the order of magnitude of the core radius. This means that the P waves are ''dead'' before they can ''tell'' the shell that the core is cylindrical, a fact which leads to a dispersion relation similar to the one of an empty shell for higher frequencies.

Figure 15 shows the complex dispersion relation of a shell filled with a shear dissipative core. The dot in Fig. 15 (mode 9) indicates the mode and the frequency for which the stress field is calculated, as represented in Figs. 21–24. Difficulties arose while calculating mode 0, in which the root-finding method jumped into mode 1. Mode 0 is therefore not visible in Figs. 15 and 16.

The dissipative nature of the core material and the fact that the density of the shell material is approximately ten times the density of the core material leads to a dynamic decoupling of the two waveguides in higher-frequency ranges. In terms of growing frequency, this decoupling appears first in the longitudinal shell mode which is coupled with the core solely through the shear modulus μ^*. Additionally, at higher frequencies the flexural mode of the shell, indicated by the slope-discontinuity of the first 14 modes,

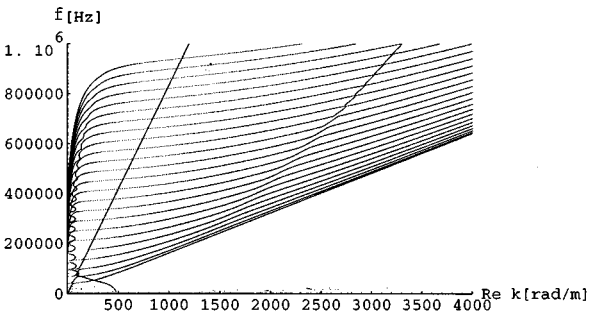


FIG. 16. Projection of the curves shown in Fig. 15 into the real plane.

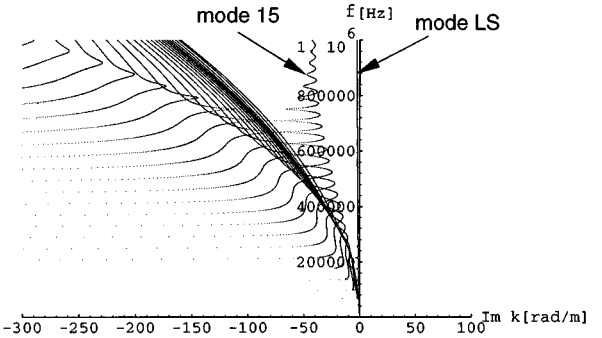


FIG. 18. Projection of the curves shown in Fig. 15 into the imaginary plane.

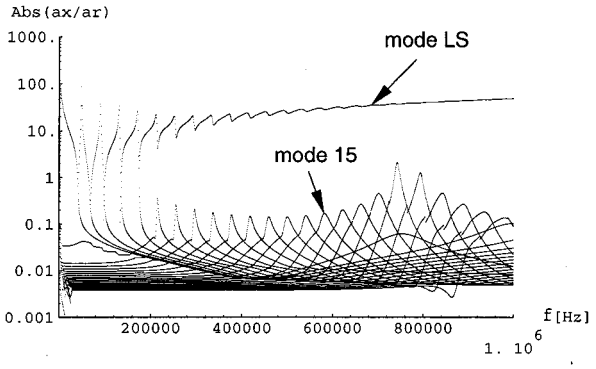


FIG. 19. Relation between axial and radial displacement of axisymmetric wave modes in a cylindrical shell containing a shear dissipative core material.

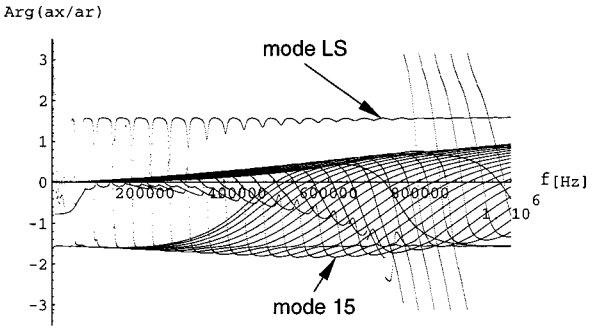


FIG. 20. Phase angle between a_x and a_r for axisymmetric modes in a shell filled with shear dissipative material according to Eq. (50).

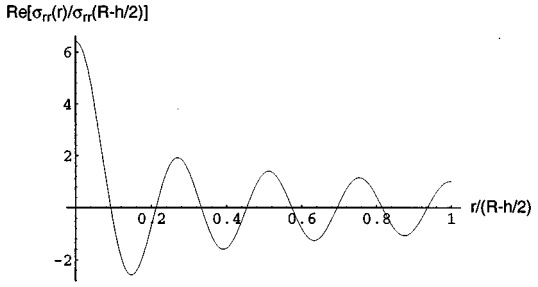


FIG. 21. Normalized radial stress σ_{rr} versus normalized radius for mode 9 at 360 kHz in a shear dissipative core.

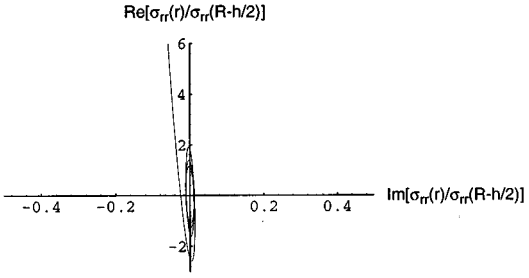


FIG. 22. Real part of the normalized radial stress σ_{rr} versus imaginary part of the normalized radial stress σ_{rr} for mode 9 at 360 kHz in a shear dissipative core.

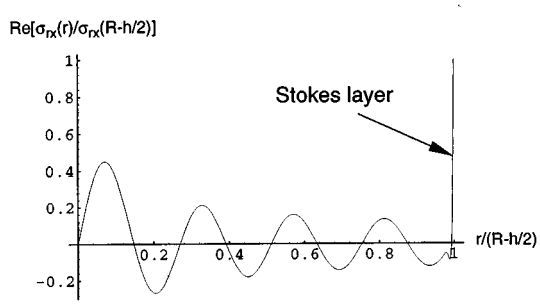


FIG. 23. Normalized shear stress σ_{rx} versus normalized radius for mode 9 at 360 kHz in a shear dissipative core.

arises as a detached mode (mode 15). The two modes in which the decoupling occurs (mode 15 and mode LS) are drawn individually in Fig. 17. The projection of the dispersion curves into the imaginary plane (Fig. 18) shows that the detached longitudinal mode (mode LS) has the lowest damping, followed by the detached flexural mode (mode 15). In Figs. 19 and 20, the absolute value and the argument of the complex displacement ratio are presented.

The following diagrams show the normalized radial stress σ_{rr} and the normalized shear stress σ_{rx} in the core material. Mode 9 and a frequency of 360 kHz are chosen as an example. In Figs. 21–23, it can be seen that the radial stress becomes a maximum on the symmetry axis, whereas the shear stress vanishes. Figure 23 shows a quickly attenuating shear boundary layer (Stokes layer) followed by a curve with increasing amplitude towards the center. This part of the diagram is governed by the radially propagating P waves. To fulfill the dynamic equilibrium conditions, the shear stress in the core σ_{rx} grows with the rate of the normal stress σ_{rr}.

The interpretation of Figs. 22 and 24 demands further explanation: In general, the normalized stress functions according to Eqs. (52) and (53) are complex and can therefore be written as follows [Eq. (56)]:

$$\tilde{\sigma}^f(\tilde{r}) = S(\tilde{r}) e^{i\alpha(\tilde{r})}, \tag{56}$$

in which a growing α means that waves are propagating from the inner surface of the shell through the core towards the center. In this case, the graphs of the real stress versus the imaginary stress appear to be more or less circle shaped. On the other hand, if α remains constant, the shell ''is riding on

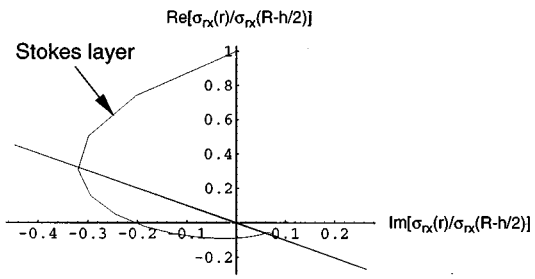


FIG. 24. Real part of the normalized shear stress σ_{rx} versus imaginary part of the normalized shear stress σ_{rx} for mode 9 at 360 kHz in a shear dissipative core.

TABLE I. Normalized parameters for three classes of core material.

Normalized core parameter	Inviscid core	Dissipative core	Shear elastic core
λ_0/E_s	4.9000×10^{-3}	4.8840×10^{-3}	4.4003×10^{-3}
$T_\lambda f_{\max}$	1.0000×10^{6}	1.0000×10^{6}	1.0000×10^{6}
μ_0/G_s	1.2800×10^{-11}	5.1200×10^{-3}	6.4000×10^{-4}
$T_\mu f_{\max}$	1.0000×10^{-1}	1.0000×10^{-2}	4.0000×10^{1}
ρ_f/ρ_s	1.2400×10^{-1}	1.2400×10^{-1}	1.2400×10^{-1}

a steady-state core vibration'' and the graph appears to be a more or less straight line. In Fig. 24 both phenomena are visible; the Stokes layer close to the shell wall can be identified as an attenuating, propagating wave, whereas the rest of the graph represents a standing wave, i.e., a vibration of the core in radial direction.

D. The shell filled with a shear elastic core

The shear elastic parameter set (see Table I) produces the most exciting results. Careful numerical analysis was carried out to ascertain that the effects presented in this section have real physical reasons and are not caused by insufficient accuracy of the root-finding procedure. In anticipation of Part II of this paper (Experimental results versus theory),[15] it should be mentioned that the curves obtained by this parameter set come closest to the measured dispersion curves, in shape and magnitude, for a shell filled with silicone oil.

In Fig. 25, modes 0, 2, 3, 5, and 6 are not plotted. The dot indicates that the stress field is calculated for mode 9 at 360 kHz. As shown in Fig. 25, increasing the elastic shear resistance leads to dispersion curves which leave the real plane periodically. Therefore the diagrams for the displacement ratio become very confusing (Figs. 26 and 27). Since the propagation velocity of p waves c_φ is almost equal for all

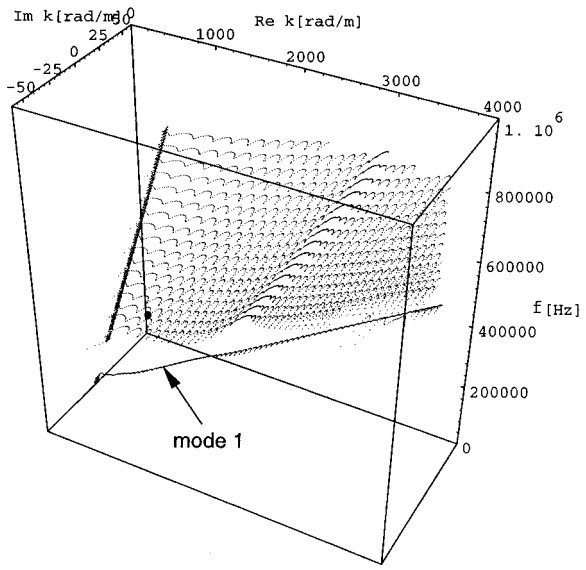


FIG. 25. Dispersion curves of axisymmetric modes of a shell filled with a shear elastic core material.

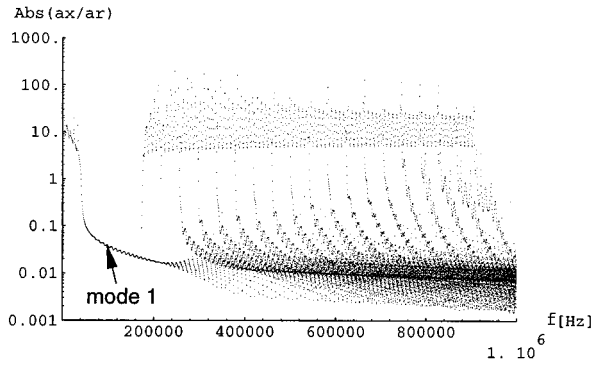


FIG. 26. Relation between axial and radial displacement of all axisymmetric wave modes presented in Fig. 25.

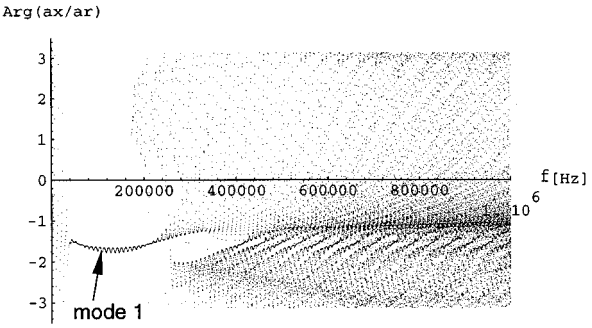


FIG. 27. Phase angle between a_x and a_r of all axisymmetric wave modes presented in Fig. 25 according to Eq. (50).

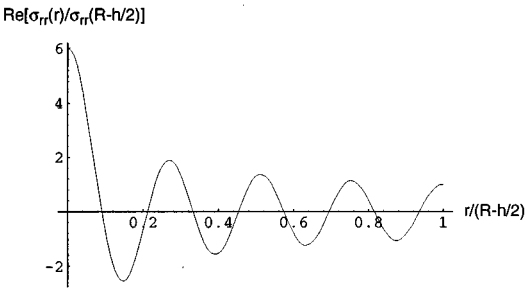


FIG. 28. Normalized radial stress σ_{rr} versus normalized radius for mode 9 at 360 kHz in a shear elastic core.

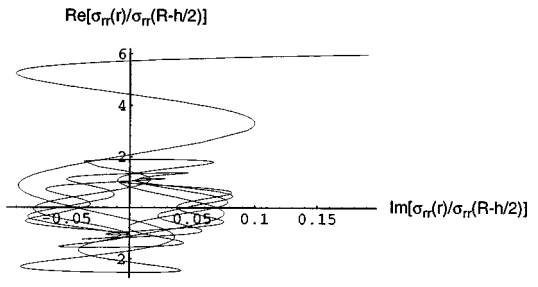


FIG. 29. Normalized shear stress σ_{rx} versus normalized radius for mode 9 at 360 kHz in a shear elastic core.

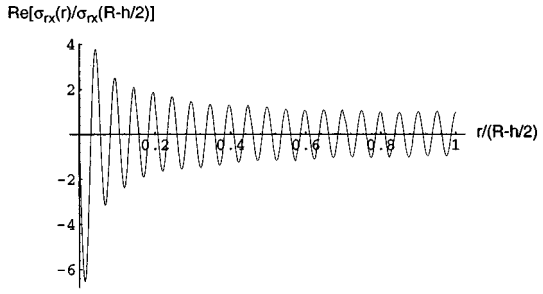


FIG. 30. Real part of the normalized radial stress σ_{rr} versus imaginary part of the normalized radial stress σ_{rr} for mode 9 at 360 kHz in a shear elastic core.

three classes of core material, the radial stress distribution presented in Fig. 28 is very similar to the one presented in Fig. 21.

As expected, major differences appear in the diagrams representing the shear behavior. When comparing Fig. 30 with Fig. 23, one can see that rather than the thin Stokes layer disappearing, shear waves propagate through the whole core with increasing amplitude towards the symmetry axis, focusing the energy.

Considering the scale of the imaginary axis of Figs. 29 and 31, one can see that the graph of the shear stress (Fig. 31) is much more ''circle shaped'' than the graph of the normal stress (Fig. 29), thus indicating a wave propagation towards the center.

E. Numbering of the modes

The numbering of the modes requires several comments. Mode 0 of the filled shell (see Fig. 10) is governed by the speed of sound in the core material and has no cutoff frequency. Mode 1 has no cutoff frequency either. Mode 2 is the first mode having a cutoff frequency, which means that one P wavelength equals the diameter of the core or the radius equals approximately a half wave length. Thus eight nodes, i.e., approximately eight half P wavelengths, can be recognized on the diagrams presenting the radial stress σ_{rr} of mode 9 at 360 kHz (Figs. 21 and 28).

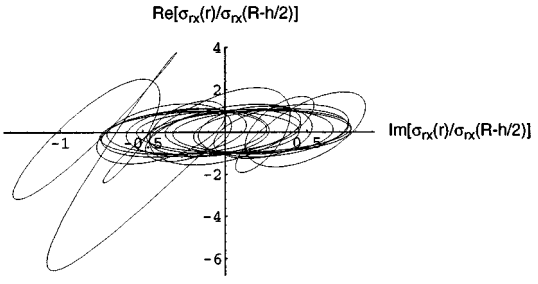


FIG. 31. Real part of the normalized shear stress σ_{rx} versus imaginary part of the normalized shear stress σ_{rx} for mode 9 at 360 kHz in a shear elastic core.

IV. CONCLUSIONS

In combining an elastic and a viscoelastic waveguide, some aspects of their interaction have been studied. Since the main interest lies in the dynamic behavior of the viscoelastic core, this investigation can be regarded as a first step on the way towards a measuring instrument for viscoelastic material properties, even though the inverse problem can be solved presently in an iterative way only.

A theory for thin-walled elastic shells is applied, which provides good results as long as the wall thickness is less than one-third of the shortest bending wavelength in the shell wall. In this investigation the shortest wavelength equals 3.2 times the wall thickness. Higher sensitivity of the shell waves to the core properties could be achieved by using a shell material with lower density. It is difficult, however, to obtain thin-walled elastic tubes of high geometrical accuracy therefore steel tubes have been used.

To check the validity of the theoretical model as well as the accuracy of the computer program, especially in the range of strong shear elasticity, the dispersion curves of axisymmetric waves in an aluminum rod with circular cross section are also measured and compared with numerical results obtained by running the same computer program as for the fluid-filled shell. Excellent agreement is found, as presented in a companion paper.[15]

ACKNOWLEDGMENT

Many thanks to Dr. Stephan Kaufmann for his kind support in numerically solving the dispersion relation.

APPENDIX: MATRIX m_{ij}

Here the coefficients of the linear system used in Eq. (44) are presented. These expressions are derived from Eqs. (37), (14), and (15). The constants D (bending stiffness of a plate) and C (longitudinal stiffness of plate) are written in terms of c_p, h, and ρ_s, because these terms can be measured more readily on an empty shell.

$$D=\frac{c_p^2 h^3 \rho_s}{12}, \quad C=c_p^2 \rho_s h, \quad \left(c_p^2=\frac{E}{\rho_s(1-\nu^2)}\right),$$

$$c_s^2=\frac{G}{\rho_s}=\frac{E}{2\rho_s(1+\nu)}, \tag{A1}$$

$$m_{11}=0, \quad m_{21}=1, \quad m_{12}=1,$$

$$m_{22}=-\frac{i\omega}{c}\left(1-\frac{c^2}{c_2^2\kappa}\right)\frac{h}{2}\cdot\frac{1}{2}\left(\operatorname{sign}\left[1-\frac{(\mathrm{Re}[c])^2}{c_2^2\kappa}\right]+1\right),$$

$$m_{13}=\beta_\varphi^* J_1(\beta_\varphi^* R), \quad m_{23}=ik^* J_0(\beta_\varphi^* R),$$

$$m_{14}=-ik^* J_1(\beta_\Psi^* R), \quad m_{24}=-\beta_\Psi^* J_0(\beta_\Psi^* R),$$

$$m_{31}=-\frac{ik^* c_p^2 h \rho_s \nu}{\kappa G R}\left(\frac{k^{*2} c_p^2 h^2 \rho_s}{12}+\kappa G-\frac{\rho_s h^2 \omega^2}{12}\right),$$

$$m_{32}=\frac{k^{*4}c_p^2h^3\rho_s}{12}-\left(\frac{c_p^2\rho_s}{\kappa G}+\frac{1}{12}\right)k^{*2}\omega^2\rho_s h^3+\frac{k^{*2}c_p^4h^3\rho_s^2}{12\kappa GR^2}$$
$$-\left(1+\frac{c_p^2\rho_s h^2}{12\kappa GR^2}\right)\rho_s h\omega^2+\frac{c_p^2h\rho_s}{R^2}+\frac{\omega^4\rho_s^2h^3}{12\kappa G},$$

$$m_{33}=\frac{-1}{12\kappa G}\,(12\kappa G-\rho_s h^2\omega^2+k^{*2}c_p^2h^2\rho_s)\Bigg\{J_0(\beta_\varphi^* R)$$
$$\times[k^{*2}\lambda^*+(\lambda^*+2\mu^*)\beta_\varphi^{*2}]-J_1(\beta_\varphi^* R)\left[\frac{2\mu^*\beta_\varphi^*}{R}\right]\Bigg\}$$
$$+(ik^*\mu^*h)J_1(\beta_\varphi^* R)[ik^*\beta_\varphi^*],$$

$$m_{34}=\frac{-1}{12\kappa G}\,(12\kappa G-\rho_s h^2\omega^2+k^{*2}c_p^2h^2\rho_s)\Bigg\{-J_0(\beta_\Psi^* R)$$
$$\times[2ik^*\mu^*\beta_\Psi^*]+J_1(\beta_\Psi^* R)\left[\frac{2ik^*\mu^*}{R}\right]\Bigg\}$$
$$+\frac{(ik^*\mu^*h)}{2}\,J_1(\beta_\Psi^* R)[k^{*2}-\beta_\Psi^{*2}],$$

$$m_{41}=\omega^2-c_p^2k^{*2},\quad m_{43}=-\frac{i\mu^*}{\rho_s h}\,J_1(\beta_\varphi^* R)[2k^*\beta_\varphi^*],$$

$$m_{42}=\frac{-ik^*\nu c_p^2}{R},\quad m_{44}=-\frac{\mu^*}{\rho_s h}\,J_1(\beta_\Psi^* R)[k^{*2}-\beta_\Psi^{*2}],\tag{A2}$$

in which $\mathrm{sgn}[x]=1$ for $x>0$, $\mathrm{sgn}[x]=-1$ for $x<0$, and $\mathrm{sgn}[x]=0$ $x=0$. $\mathrm{Re}[c]$ takes the real part of the complex phase velocity c.

[1] L. Pochhammer, ''Ueber die Fortpflanzungsgeschwindigkeit kleiner Schwingungen in einem unbegrenzten isotropen Kreiscylinder,'' Journal für die reine und angewandte Mathematik. In zwanglosen Heften (Berlin) **81**, 324–336 (1876).

[2] N. Joukowsky, ''Ueber den hydraulischen Stoss in Wasserleitungsrohren,'' Mémoir de l'Academie Impérial de Sciences de St. Petersburg, Ser. 8 **9**(5), 5 (1990).

[3] R. Skalak, ''An Extension of the Theory of Water Hammer,'' Paper No. 55-S-18, Hydraulics Division, ASME, Baltimore, MD, 1955.

[4] C. R. Fuller and F. J. Fahy, ''Characteristics of wave propagation and energy distributions in cylindrical elastic shells filled with fluid,'' J. Sound Vib. **81**, 501–518 (1982).

[5] T. C. Lin and G. W. Morgan, ''Wave propagation through fluid contained in a cylindrical, elastic shell,'' J. Acoust. Soc. Am. **28**, 1165–1176 (1956).

[6] Y. P. Guo, ''Approximate solutions of the dispersion equation for fluid-loaded cylindrical shells,'' J. Acoust. Soc. Am. **95**, 1435–1440 (1994).

[7] Y. P. Guo, ''Attenuation of helical wave radiation from cylindrical shells by viscoelastic layer,'' J. Acoust. Soc. Am. **97**, 298–308 (1995).

[8] T. J. Plona, B. K. Sinha, S. Kostek, and S.-K. Chang, ''Axisymmetric wave propagation in fluid-loaded cylindrical shells. II: Theory versus experiment,'' J. Acoust. Soc. Am. **92**, 1144–1155 (1992).

[9] B. K. Sinha, T. J. Plona, S. Kostek, and S.-K. Chang, ''Axisymmetric wave propagation in fluid-loaded cylindrical shells. I: Theory, '' J. Acoust. Soc. Am. **92**, 1132–1143 (1992).

[10] L. D. Lafleur and F. D. Shields, ''Low-Frequency propagation modes in a liquid-filled elastic tube waveguide,'' J. Acoust. Soc. Am. **97**, 1435–1445 (1995).

[11] S. P. Timoshenko, ''On the Correction for Shear of the Differential Equation for Transverse Vibrations of Prismatic Bars,'' Philos. Mag. **41**, 125–131 (1921).

[12] G. R. Cowper, ''The Shear Coefficient in Timoshenko's Beam Theory,'' J. Appl. Mech. **33**, 335–340 (1966).

[13] G. Herrmann and I. Mirsky, ''Three-Dimensional and Shell-Theory Analysis of Axially Symmetric Motions of Cylinders,'' J. Appl. Mech. **23**, 563–568 (1956).

[14] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1968).

[15] J. Vollmann, R. Breu, and J. Dual, ''High-resolution analysis of the complex wave spectrum in a cylindrical shell containing a viscoelastic medium. Part II. Experimental results versus theory,'' J. Acoust. Soc. Am. **102**, 909–920 (1997).

High-resolution analysis of the complex wave spectrum in a cylindrical shell containing a viscoelastic medium. Part II. Experimental results versus theory

Johannes Vollmann,[a] Roger Breu, and Jürg Dual
Institute of Mechanics, ETH Zürich, Swiss Federal Institute of Technology, CH-8092 Zürich, Switzerland

(Received 19 December 1996; accepted for publication 7 April 1997)

The complex frequency spectrum of axisymmetric wave modes in a circular cylindrical shell containing various viscoelastic media is measured. A new measurement technique has been developed for this purpose by combining a high-resolution laser interferometer with modern spectrum estimation methods. To decompose the complex wave-number dependence, a complex spectrum estimation method has been implemented. Up to 40 dispersion curves of traveling, axisymmetric modes are decomposed simultaneously in a frequency range between 1 kHz and 2 MHz. The guided structural waves are excited by piezoelectric transducers. Linear elasticity can be considered as an extreme case of viscoelasticity (long relaxation times compared with the deformation periods). To ascertain the validity of the theory, dispersion curves are calculated for a shell containing a viscoelastic material behaving like the elastic shell and are compared with the measured curves of an isotropic aluminium rod. The phenomenon of ''backward wave propagation,'' in which the group velocity and the phase velocity of one mode have opposite signs, is clearly measured. Excellent agreement between experimental and theoretical results, which are also presented in a corresponding paper [J. Vollmann and J. Dual, J. Acoust. Soc. Am. **102**, 896–908 (1997)], is found over a wide parameter range, including the case of a linear elastic rod.
 [S0001-4966(97)01208-3]

PACS numbers: 43.35.Mr, 43.40.Ey, 43.40.Cw, 43.20.Tb [HEB]

INTRODUCTION

The propagation of structural waves in shells filled or surrounded with liquid is of fundamental interest in various applications, ranging from nondestructive evaluation (of the shell, a layer, or the liquid) to scattering phenomena caused by submerged, purely elastically and viscoelastically damped shells interacting with waves in the surrounding fluid. A historical overview discussing previous work in the field of fluid-loaded cylindrical shells is delineated by Fuller and Fahy[1] and in the corresponding theoretical part of this investigation.

In 1977 Krause *et al.*[2] recorded and published some plots of the strain versus time in a fluid-filled tube which was subjected to a longitudinal impact. In addition these results were compared with those of the extended Skalak water hammer theory[3] by Barez *et al.*[4] in 1979.

It is generally observed that the number of papers treating this subject experimentally is rather limited when compared with the various theoretical publications which differ mainly in their objectives and assumptions. This might be due to the fact that the number of traveling modes grows with frequency and the group and phase velocities of the various modes are almost equal over wide parameter ranges. Thus the dispersion relation can scarcely be detected by means of measuring arrival times of narrow-band pulses or by calculating phase differences of propagating narrow-band pulses at two locations, except for the first two modes in a very low-frequency range, as recently presented by Lafleur and Shields.[5] For a linear elastic cylindrical rod, a simple but more effective method was used by Zemanek.[6] He was able to measure up to six symmetric and antisymmetric modes in an elastic cylindrical rod with a circular cross section measuring the frequencies of standing waves.

In this paper, a new method is presented to determine the complex dispersion relation directly from a series of displacement recordings along a shell. To explain the metrology, the nature of guided waves propagating along structures of finite cross-sectional dimensions should be outlined first. Small disturbances in an elastic or viscoelastic medium produce *P* and *S* waves. Interference of these *P* and *S* waves and their reflections from the boundaries of the structure leads to ''wave modes'' propagating along the structure. These modes may differ in shape, group velocity, phase velocity, and attenuation. In the case of axisymmetric waves in a cylindrical structure, the propagating waves can be described as follows:

$$u(x,t)=\sum_{n=1}^{n=p}\left(A_n e^{i(\omega t+k_n x)}+B_n e^{i(\omega t-k_n x)}\right), \tag{1}$$

in which u denotes the displacement of a surface point of the structure and p the number of traveling modes in the x direction. ω indicates the circular frequency and k the complex wave number which can be regarded as ''frequency'' in space domain. The imaginary part of the wave number k indicates a decaying or a growing wave mode depending on its sign. A dispersion diagram contains characteristic infor-

[a]Current address: Northwestern University, Center for Quality Engineering and Failure Prevention, 2137 N. Sheridan Road, Evanston, IL 60206-3020, Electronic mail: vollmann@nwu.edu

mation about geometry, frequency-dependent material behavior, and homogeneity of a structure. It is a special advantage of axisymmetric waves in a cylindrical structure that no boundary conditions causing radiation loss have to be considered and that a three-dimensional structure can be inspected or characterized by measuring dispersion relations of one-dimensional wave propagation.

To measure a dispersion relation in a defined interval of frequency and wave number, a spectrum analysis in time and space domain is required. Applying discrete spectrum analysis methods, two features determining the number of samples and the distance between two consecutive samples (in time and space) must be observed. First, the highest harmonic component, i.e., the shortest cycle or the shortest wavelength, must to be sampled at least two points in time and space, respectively.

The second feature concerns the resolution. The question here is how many samples are necessary to separate two neighboring (in terms of frequency or wave number) modes. The resolution of discrete Fourier transformation methods such as fast Fourier transform (FFT) grows linearly with the length of the interval, i.e., the number of samples and is therefore intrinsically limited by the method. Modern spectrum estimation methods approximate given signals with a harmonic model. Moreover, highly sophisticated methods are able to isolate signals of interest from surrounding noise always present in experimental data. These methods do not have inherent resolution limits, although the prediction quality grows with the number of samples.

In this investigation, a combination of discrete spectrum analysis (FFT) and spectrum estimation (linear prediction and total least squares) is seen to be very effective. FFT is applied in order to obtain the complex frequency spectrum versus frequency of recorded shell displacements versus time. In the space domain a spectrum estimation method is used to decompose the complex wave-number dependence. Thus the number of required samples in the space domain is reduced to less than a sixth compared with FFT. An additional advantage is the fact that the wave numbers are decomposed, including their signs indicating the direction of propagation. Thus outgoing pulses can be distinguished from reflected pulses.

Spectrum estimation methods are a still-growing topic in signal processing, and improved algorithms are published frequently. Comprehensive introductions are given by Kay,[7] Marple,[8,9] and Hayes,[10] as well as in German by Oestreich.[11] A comparison of several methods based on singular value decomposition with FFT is presented by Uike *et al.*[12] The algorithm implemented in this investigation, originally developed for the signal processing of NMR signals, was published by Tirendi and Martin[13] and is based on the work of Golub and Van Loan.[14,15]

To the authors' knowledge, the work presented in this investigation is one of the first applications of spectrum estimation methods used in structural mechanics to characterize dynamic material properties.

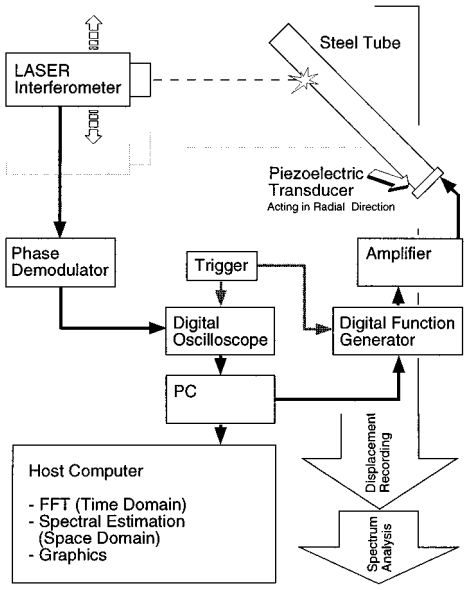


FIG. 1. Experimental setup.

I. EXPERIMENTS

Numerous combinations of shells and fluids (water, milk, alcohol, glycerine, silicone oil, and polyisobutylene) were investigated in this project. An extract of these experimental results outlining the main aspects and phenomena will be presented in this paper. The shell remains unchanged in geometry and material. A steel tube with a diameter of 26 mm, a wall thickness of 0.5 mm, and a length of 1 or 2 m is used to perform the experiments.

To verify the theoretical model derived in Part I,[16] as well as the numerical accuracy of the root-finding algorithm in the parameter range of strong shear elasticity, dispersion curves of axisymmetric waves in an isotropic aluminum rod with a diameter of 18 mm are also measured and compared with the numerical results.

A. Experimental setup

The order of magnitude of the shell displacements amounts to less than 10 nm. To be able to detect such movements and to avoid any disturbance of the traveling waves caused by the measurement itself, the shell displacement function is measured with a heterodyne laser interferometer (Sensor Head: Polytec OFV 300, Demodulator: Polytec OFV 2100) which represents the most essential part of the experimental setup shown in Fig. 1. The commercial configuration of this interferometer did not satisfy the resolution requirements over the frequency range of our interest. Therefore a demodulator for the phase-modulated signal with a 40-MHz carrier frequency was developed by Dual and Hägeli[17] at the Institute of Mechanics, ETH Zürich. This phase demodulator allows the measurement of surface displacements over a frequency range from 1 kHz to 8 MHz with a resolution of 10^{-13} m/$\sqrt{\text{Hz}}$. For a bandwidth of 1 MHz, this corresponds to a noise level of about 1 Å before averaging.

Typically the waves are excited by a ring-shaped piezoelectric transducer subjected to a linear sweep pulse with a start frequency of 400 kHz, a stop frequency of 1250 kHz,

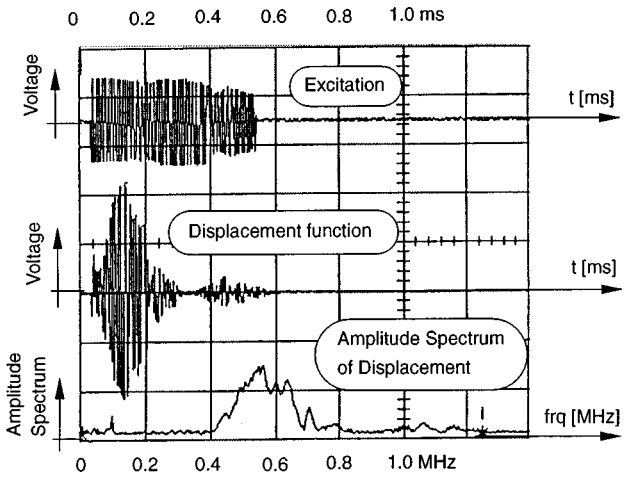


FIG. 2. Screen print of the digital oscilloscope displaying: The voltage of the excitation signal (linear sweep, 400–1250 MHz, 0.5 ms), average of 300 displacement functions (measured at one point), amplitude spectrum of the displacement function above (fluid: silicone oil), sampling rate: 5 MHz.

and a duration of 0.5 ms. This pulse is applied repeatedly with a trigger frequency of 20 Hz. The displacement of the shell surface versus time caused by the outgoing pulse and its reflections is measured and averaged 300 times before the laser interferometer is shifted to the next location. The displacement functions are recorded at up to 160 points along the tube, depending on the number of expected modes in the frequency range considered.

A typical signal is presented in the screen print of the digital oscilloscope in Fig. 2. Here the top graph shows the voltage of the excitation signal and the lower graph the averaged displacement function followed by its amplitude spectrum.

B. Piezoelectric excitation

Piezoelectric transducers are used to convert voltage into mechanical stress. Since the ceramic material (lead-zirconate-titanate) is polarized after firing, transducers are available in various shapes and with various polarization directions. Figure 3 shows the cross section of a tube and the attached piezoelectric ring transducer. The best signals are achieved when the transducer is fixed at an open end of the shell.

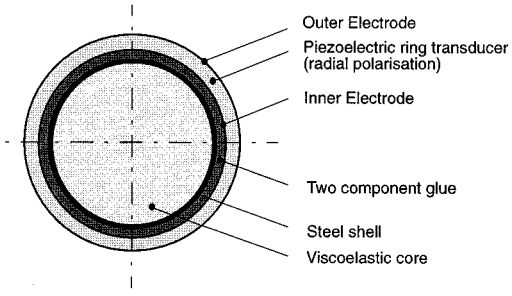


FIG. 3. Cross section of a steel tube with piezoelectric ring transducer acting in radial direction.

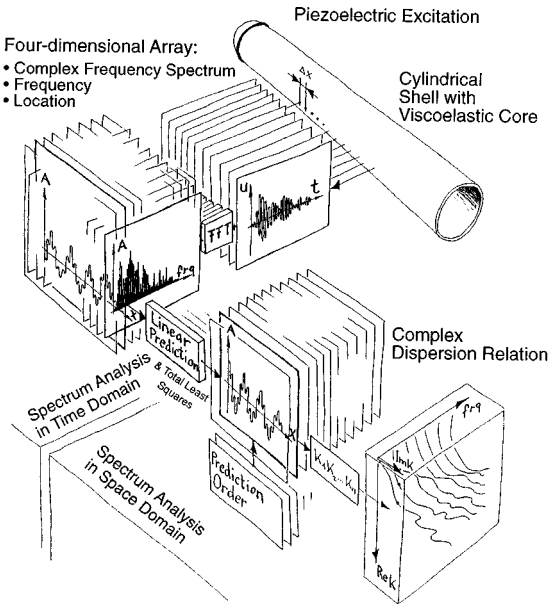


FIG. 4. Visualization of the major steps in signal processing.

II. SIGNAL PROCESSING

Figure 4 illustrates the digital data processing leading from the displacement measurement to the complex dispersion relation. The highly dispersive broadband pulse is propagating forward and backward from one end of the tube to the other until its amplitude has decayed below the noise level. The best results are obtained when the whole transient response of the shell to one pulse is recorded to avoid leakage. This demand is, however, a question of storage capacity of the oscilloscope and attenuation characteristics of the structure as well as the frequency range considered. In other words, since the storage capacity of the oscilloscope is limited, one has to balance between a short time window with a high-sampling frequency or a long window with a low-sampling frequency to achieve good measurement over a wide frequency range. Therefore moderately damped systems are more suitable for broadband frequency measurement than systems with low damping, even though the situation of those systems can be improved by windowing.

Typical values for measurements with an upper frequency limit of 1 MHz and a storage capacity of $n_t = 8192$ samples lead to a frequency resolution of approximately 250 Hz according to Eq. (2). (1-MHz signal frequency demands a sampling frequency of at least 2 MHz, thus $\Delta t_{\text{sample}} \leqslant 0.5\ \mu\text{s}$.)

$$\Delta f = \frac{1}{n_t \Delta t_{\text{sample}}} = 244\ \text{Hz}. \tag{2}$$

A. Spectrum analysis in space domain

Sampling the displacement function versus time is an easy task since it is done automatically by the digital storage oscilloscope. To sample the displacement function along the shell, the laser interferometer is moved from one point to another by a computer-controlled stepping motor. According to Eq. (3), applying FFT in the space domain would require

at least $n_s=800$ samples along the shell to achieve a wave-number resolution of 10 rad/m for a given upper wave-number limit of 4000 rad/m ($\Delta x_{\text{sample}}=0.79$ mm). A wave-number resolution of 10 rad/m represents a relative resolution which is still ten times lower than the relative resolution in the frequency range, but it is sufficient to obtain good results for the systems considered in this investigation:

$$\Delta k=\frac{2\pi}{n_s\Delta x_{\text{sample}}}=10 \text{ rad/m}. \tag{3}$$

To circumvent the onerous task of recording a minimum of 800 samples in the space domain, a spectrum estimation method is implemented rather than FFT. The general concept for these types of methods dates back to 1795 when Prony[18] presented a procedure to express an arbitrary function in terms of a linear combination of exponential functions.

The condition that the shortest wavelength of interest must be sampled at at least two points, thus determining the distance between two consecutive measurements in the space domain, remains indispensable.

As illustrated in Fig. 4, the spectrum estimation method is applied for each frequency given by the discrete Fourier transform. It consists of two major steps: First, the discrete frequency spectrum versus location (for one particular frequency) can be expressed in terms of exponential functions [see Eq. (5)]. This step is called ''linear prediction'' (LP). Since the values of the frequency spectrum are available at discrete locations only, where the amplitude versus time function is measured, the location x is described in terms z and Δx_{sample} [see Eq. (4)]:

$$x=z\Delta x_{\text{sample}}, \tag{4}$$

in which z indicates the sampling index along the shell ($z=0,1,2,\ldots,N-1$);

$$A(z)\approx C_1\mu_1^z+C_2\mu_2^z+\cdots+C_n\mu_n^z. \tag{5}$$

$A(z)$ represents the value of the frequency spectrum at the location z for one particular frequency, and $\mu_i=e^{k_i\Delta x_{\text{sample}}}$. C_i and k_i are to be determined.

The linear prediction procedure is outlined below according to Hildebrand.[19] For an autoregressive notation of a similar procedure which reaches the same goal, the reader is referred to Marple.[20]

$$\begin{aligned}
&C_1+C_2+\cdots+C_n=A(z=0),\\
&C_1\mu_1+C_2\mu_2+\cdots+C_n\mu_n=A(z=1),\\
&C_1\mu_1^2+C_2\mu_2^2+\cdots+C_n\mu_n^2=A(z=2),\\
&\cdots=\cdots,\\
&C_1\mu_1^{N-1}+C_2\mu_2^{N-1}+\cdots+C_n\mu_n^{N-1}=A(z=N-1).
\end{aligned} \tag{6}$$

Since C_i and μ_i are unknowns, at least $N=2n$ equations are needed, in principle, to solve the system presented in Eq. (6).

Further steps must be introduced to solve this highly nonlinear system. Let $\mu_1,\ldots,\mu_n$ be roots of Eq. (7):

$$\mu^n+\alpha_1\mu^{n-1}+\alpha_2\mu^{n-2}+\cdots+\alpha_{n-1}\mu+\alpha_n=0. \tag{7}$$

To determine the coefficients $\alpha_1,\ldots,\alpha_n$, the first equation of Eq. (6) is multiplied by α_n, the second by $\alpha_{n-1},\ldots$, the nth by α, and the $(n+1)$th by 1. All equations obtained in this manner are added. The fact that each μ satisfies Eq. (7) allows the result to be written in the following form [see Eq. (8)]:

$$A_n+\alpha_1A_{n-1}+\cdots+\alpha_nA_0=0, \tag{8}$$

in which $A_n=A(z=n)$. A set of $N-n-1$ additional equations can be obtained in the same way by starting successively with the second, third,..., $(N-n)$th equation leading to the linear system shown in the following equation;

$$\begin{aligned}
&A_n+A_{n-1}\alpha_1+A_{n-2}\alpha_2+\cdots+A_0\alpha_n=0,\\
&A_{n+1}+A_n\alpha_1+A_{n-1}\alpha_2+\cdots+A_1\alpha_n=0,\\
&\cdots=\cdots,\\
&A_{N-1}+A_{N-2}\alpha_1+A_{N-3}\alpha_2+\cdots+A_{N-n-1}\alpha_n=0.
\end{aligned} \tag{9}$$

Equation (9) can be written in matrix form:

$$A\boldsymbol{\alpha}=\mathbf{b}. \tag{10}$$

For $N=2n$, the matrix A has full rank and the linear system [Eq. (10)] has a unique solution α, which can be directly calculated by inversion of the matrix A. The values of μ_i are then obtained as roots of Eq. (7). With the known values of μ_i, the system of Eq. (6) becomes linear, which allows C_i to be determined. The ''beauty'' of this method introduced by Prony lies in the fact that the nonlinearity of the problem, which is caused by the exponential approximation, is concentrated in Eq. (7), which can be easily solved numerically. The procedure described up to this point is called ''forward linear prediction'' (FLP).

Provided that the values of A_i consist of exponential components caused by propagating waves only, the problem would already be solved by the procedure described above. The number of required samples N would then be two times the number of expected modes m, since C_i and μ_i must be determined and the number of expected modes m would equal the polynomial order n. In reality, however, the situation is more complicated, demanding a second major step. The signals of interest are hidden in noise and many more samples than $2n$ are needed. Moreover, various nonaxisymmetric wave modes are also excited and therefore the number of expected modes m (prediction order) is not known *a priori*. In this context, some questions arise. What polynomial order n, as a function of the number of samples N, should be chosen for the polynomial introduced in Eq. (5)?

Two cases shall be discussed:

(i) $n=m$ and $N>2n$. In this case, the linear system introduced in Eq. (10) is overdetermined and can be solved by standard least-squares methods. This approach has no spare parameters to ''describe'' the noise and is therefore not providing optimal results by means of distinguishing harmonic components from surrounding noise.

(ii) $n>m$ and $N>2m$. Here the linear system of Eq. (10) might become underdetermined and an infinite number of solutions exist. The pseudoinverse $A^{\#}$ [see Eq.

(11)] of the matrix A provides a unique solution having minimum norm, which means that the sum of the squares of all entries in $AA^{\#}-I$, in which I is an identity matrix, is minimized;

$$\boldsymbol{\alpha}=A^{\#}\mathbf{b}. \tag{11}$$

For the signals analyzed in this investigation, a polynomial order n, which is higher than the number of expected modes m, is seen to provide better results. Tufts and Kumaresan[21] suggest a polynomial order of $n=3N/4$ to achieve best accuracy in the determination of the wave numbers. This result was obtained by analyzing synthetic data consisting of two exponential functions and noise. It is also verified by the authors of this investigation for four signal components (sinusoids) with noise and is thus adapted for the signal processing of this investigation. The authors would like to remark, however, that the optimal polynomial order could depend on the actual signal-to-noise ratio configuration and therefore demands further investigation. At this point the measured signals consisting of the harmonic components of interest and the noise are described by a parametric model. The next step is the separation of the harmonic components from the noise, a task which is integrated into the calculation process of the pseudoinverse of A.

By applying singular value decomposition, the matrix A can be written in the following form [see Eq. (12)]:

$$A=U^{H}\Sigma V, \tag{12}$$

in which U^{H} denotes the Hermitian transpose of the matrix U. (The Hermitian transpose, x^{H} is the complex conjugate of the transpose of x.) Both U and V are row orthonormal matrices. Σ is a diagonal matrix containing the singular values σ_i of the matrix A. The pseudoinverse of A is defined in terms of U, Σ, and V as follows [see Eq. (13)]:

$$A^{\#}=V^{T}\Sigma^{-1}U. \tag{13}$$

But before this pseudoinversion is carried out, the singular values σ_i are monitored. High values of σ_i are associated with the harmonic components of interest, whereas ''small'' values of σ_i are associated with noise and are therefore set to zero. The corresponding columns in V and U are also set to zero, leading to V_q and U_q. In order to realize the quality of the series of measurements, the decision as to what a ''small'' singular value means is made in an interactive way for all measurements presented in this investigation, thus determining the prediction order m. With the m remaining singular values, a new matrix $\hat{\Sigma}$ containing ''less noise'' is rebuilt:

$$\hat{A}^{\#}=V_q^{T}\hat{\Sigma}^{-1}U_q. \tag{14}$$

And finally, Eq. (11) is solved with the improved pseudoinverse $\hat{A}^{\#}$.

The procedure described up to this point is called ''singular value decomposition (SVD) of the data matrix A.'' It already provides very good results when applied to the decomposition of the wave numbers from the data sets treated in this dissertation. Further improvements are achieved by applying the method of ''total least squares'' (TLS). The main ideas of this method are outlined below: In a classical linear least-squares problem, as presented in Eq. (15), a linear mathematical model is intended to fit to given (measured) data $\mathbf{b}+\boldsymbol{\epsilon}$:

$$A\boldsymbol{\alpha}=\mathbf{b}+\boldsymbol{\epsilon}. \tag{15}$$

Thus the random errors $\boldsymbol{\epsilon}$ are confined to the right-hand side, the ''observation side,'' of Eq. (15). The problem of an exponential approximation is not a linear least-squares problem, but it is reduced to one by the linear prediction procedure. Therefore Eq. (10) cannot be discussed in terms of a ''mathematical model'' on the left-hand side and an ''observation vector'' on the right-hand side. When analyzing real measurements, both sides contain ''model'' and ''noisy observation'' features and a more appropriate version of Eq. (10) should be written as shown in Eq. (16).

$$(A+E)\boldsymbol{\alpha}=\mathbf{b}+\boldsymbol{\epsilon}, \tag{16}$$

in which E and $\boldsymbol{\epsilon}$ represent random errors (noise).
The total least-squares method provides a solution $\boldsymbol{\alpha}$ of Eq. (16) while fulfilling the following condition [Eq. (17)] in order to minimize influence of the errors on $\boldsymbol{\alpha}$:

$$\text{minimize}_{\boldsymbol{\epsilon},E}\|[\boldsymbol{\epsilon},E]\|_F, \tag{17}$$

in which $[\boldsymbol{\epsilon},E]$ denotes an augmented matrix (Signifies a matrix the elements of which are the coefficients of a set of simultaneous linear equations with the terms of the equations entered in an added column.) $\|\cdot\|_F$ denotes the Frobenius norm [Frobenius, Ferdinand Georg, 1849–1917. German mathematician. Professor at the Eidgenössische Polytechnikum, Zürich (1875) and at the Univ. of Berlin (1892).] The Frobenius norm of a $m\times n$ matrix M is defined by Eq. (18):

$$\|M\|_F=\sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}m_{ij}^2}. \tag{18}$$

If a minimizing pair $[\boldsymbol{\epsilon},E]$ is found, then any $\boldsymbol{\alpha}$ satisfying Eq. (16) is said to solve the total least-squares problem. The total least-squares problem has been analyzed in terms of the singular value decomposition by Golub and Van Loan[14] in 1980.

To evaluate the prediction order, singular value decomposition is applied to the augmented matrix C:

$$C=[\mathbf{b},A]=U\Sigma V^{H}. \tag{19}$$

To be entirely correct, the evaluation of the prediction order m (number of singular values of C which are considered to carry information about the harmonic components) should be made for all frequencies given by the FFT (see Fig. 4). But since the FFT provides thousands of samples, the prediction order is evaluated in a preprocess for a selection of frequencies only which are distributed over the whole frequency range by applying singular value decomposition of C in an interactive way. Since the polynomial order is kept constant at $n=3N/4$, a wrong prediction order scarcely affects the accuracy of the decomposition, but it leads to either modes being omitted or noise being interpreted as a harmonic component, which results in an additional dot on the dispersion diagram.

Then the whole frequency range is subdivided into coherent intervals of constant prediction order (within an interval) to be finally decomposed for each frequency. Therefore the vector $\boldsymbol{\alpha}$ [see Eq. (16)] is calculated according to Tirendi and Martin.[13]

A complete understanding of this procedure would demand several further steps; however, for briefness, the results only are presented below. From V, the matrix V_2 is formed out of the $(m+1)$ th to the $(n+1)$ th column v_i [see Eq. (20)]:

$$V_2=[v_{m+1}\cdots v_{n+1}]. \tag{20}$$

Here, $(v_j)_1$ is the jth element in the first row of V_2 and v_j' is the vector formed by the remaining elements of the jth column [see Eq. (21)]:

$$v_j=\begin{bmatrix}(v_j)_1\\ v_j'\end{bmatrix}. \tag{21}$$

The vector $\boldsymbol{\alpha}$ can now be calculated as presented in Eq. (22):

$$\boldsymbol{\alpha}=\sum_{j=m+1}^{n+1}\left[\frac{(v_j)_1^*}{\sum_{i=m+1}^{n+1}|(v_i)_1|^2}\right]v_j', \tag{22}$$

in which $(v_j)_1^*$ represents the complex conjugate of $(v_j)_1$. Now the coefficients $\alpha_1,\ldots,\alpha_n$ are known and the values μ_p are the roots of Eq. (7). With the values μ_i, Eq. (6) becomes a linear system for C_i and the complex wave numbers k_i can be determined.

B. Confidence and accuracy

A two-dimensional spectrum analysis consisting of FFT for the decomposition of the frequencies; and a combination of linear prediction and total least squares for the decomposition of the wave numbers of propagating wave modes, enables the simultaneous measurement of a considerable number of dispersion curves in the space defined by frequency and complex wave number. The following questions concerning the reliability and accuracy of this new method may arise:

(1) Is the method free of systematic errors which reach a level that they could affect the accuracy of the determination of material properties? In statistical terms, is the method biased (bias: systematic deviation of a statistical estimate from the quantity it estimates), unbiased, or asymptotically unbiased?
(2) How do random errors of the input data propagate through the decomposition process?
(3) What is the lowest signal-to-noise ratio (SNR), the so-called ''threshold,'' at which the method provides reasonable results?
(4) How does the prediction order m affect the accuracy of the decomposed signals?

Since spectral estimation methods are new, and are applied in this investigation particularly to parameter ranges where other methods failed to work, some of the questions above cannot yet be answered on a quantitative basis and require further investigation.

Evidence that the method appears to be free of systematic errors, however, is given by dispersion curves which are governed by one or two material properties which are very well-known, as presented in the following section. In other words, wherever the method could be verified, it showed perfect agreement with other methods and therefore appears to be unbiased.

Each spectrum analysis (in time or space domain) is an independent process (see Fig. 4), and random errors of the input quantities and the process itself are therefore directly visible in the dispersion diagram. The better the dots of the single measurements ''line up'' to a continuous graph, the lower the influence of random errors.

For a detailed analysis of the total least-squares method, especially of its statistical and sensitivity properties, the reader is referred to the work of Sabine Van Huffel.[22]
Another feature, quantifying the quality of a spectrum estimation method, is the lowest signal-to-noise ratio (threshold) for which the method provides reasonable results. The signal-to-noise ratio is defined as follows [see Eq. (23)]:

$$\mathrm{SNR}=10\log_{10}\left(\frac{a^2}{2\sigma_n^2}\right), \tag{23}$$

in which a represents the amplitude of the signal and σ the variance of the complex Gaussian white noise. The method applied in this investigation (linear prediction and total least squares) provides reasonable results down to a SNR level of 7 dB, which means that the amplitude of the signals amounts to 3.17 times the variance of the surrounding noise. A numerical analysis of the accuracy of several methods, based on singular value decomposition as a function of the SNR, is given by Uike *et al.*[12]

An analysis of the influence of the polynomial order n and the prediction order m on the detectability and the accuracy of the signals is given by Tufts and Kumaresan.[21]

III. EXPERIMENTAL RESULTS IN COMPARISON WITH THEORETICAL RESULTS

To keep this report comprehensible, only a selection of results elucidating the influence of the core material on the topology of the dispersion diagram is presented below. All diagrams refer to a steel shell with a diameter of 26 mm, a wall thickness of 0.5 mm, and a length of 2 m, or to an isotropic aluminum rod with a circular cross section and a diameter of 18 mm.

A. The empty shell

The properties of the shell material are determined from the longitudinal mode in an empty shell which is excited with a longitudinally acting piezo transducer at one end. Figure 5 shows the measured dispersion curves of an empty steel shell. The best fit for the longitudinal mode and for the coupling area is achieved with the following parameters: Young's modulus $E=2.000\times10^{11}$ N/m^2, Poisson ration $\nu=0.28$, density $\rho=7800$ kg/m^3. A value of $\kappa=0.8$ was chosen from Cowper[23] for the shear correction factor. The corresponding theoretical model is presented in Part I.[16] In Fig. 6 the experimental and the theoretical curves are compared,

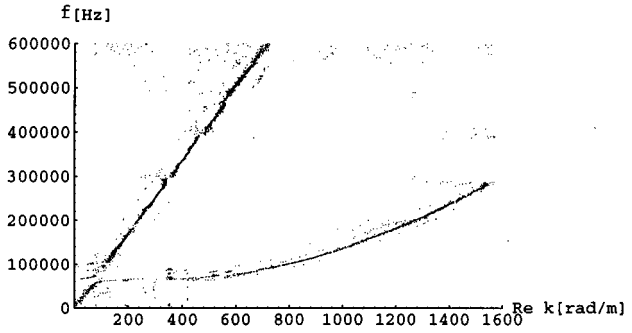


FIG. 5. Measured dispersion curves of an empty shell providing the values for the shell material. Number of samples in space domain: $n_s=80$; $\Delta x_{\text{sample}}=2.0$ mm.

showing good agreement for the longitudinal mode and for the coupling area. The slight difference for the flexural mode is probably caused either by a boundary layer of air diminuating its wave speed, an effect which is not considered in this calculation, or by a nonoptimal value for κ.

B. The shell filled with silicone oil in the low-frequency range

One of the first series of measurements was performed on a silicone oil (Dow Corning® 200 Fluid) with a density of 969.0 kg/m^3 and a nominal viscosity of 60 000 mm^2/s which means that the viscosity equals roughly 60 000 times the value of water at 20 °C. The primary question was whether the dynamic behavior of the silicone core could be at all described by the viscoelastic models introduced in Part I[16] and if so, whether it could be classified (more or less) as an inviscid, a dissipative, or a shear elastic core. At the time these curves were measured, a commercial phase demodulator, limiting the frequency range to about 300 kHz, was used.

In reference to the diagram in Fig. 7 and its projection into the real plane in Fig. 8, two features should be emphasized. The curves in the half space of positive real wave numbers represent outgoing waves from the piezoelectric excitation. The strong damping caused by the highly viscous oil leads to the fact that little information is received from the reflected waves propagating towards the excitation. The steep parts of the dispersion curves are associated with the longitudinal motion of the shell. Since the density of the

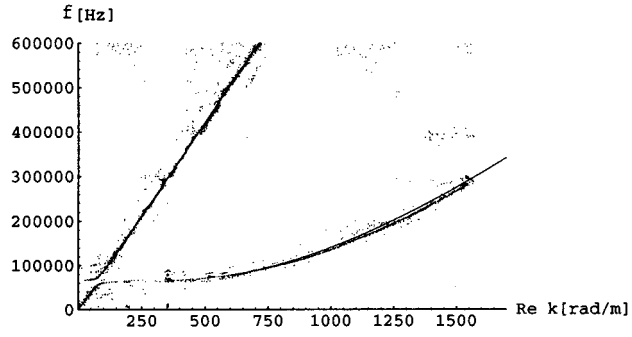


FIG. 6. Measured dispersion curves versus calculated curves of an empty shell.

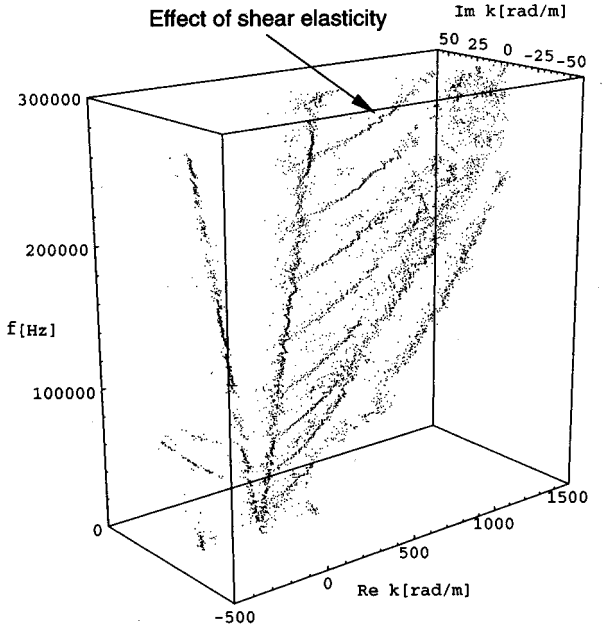


FIG. 7. Measured dispersion curves of the first nine axisymmetric wave modes in a cylindrical shell filled with silicone oil with a nominal viscosity of 60 000 mm^2/s; temperature: 24.1 °C. Number of samples in space domain: $n_s=120$; $\Delta x_{\text{sample}}=2.0$ mm.

shell equals approximately eight times the density of the fluid, the interaction is relatively weak along these parts of the curves, and reflected signals can therefore be clearly identified.

Another feature, regarding the two highest modes (indicated by the arrow in Fig. 7) can be observed. Here the curves tend to leave the plane defined by frequency and real wave number periodically. Based on Part I,[16] such behavior is also visible in theoretical results when the shear elasticity of the viscoelastic core is increased. This effect is small, however, when compared to the sensitivity of the system to changes in the velocity of P waves in the core. Therefore the experimental data of Fig. 8 were used to determine a first set of parameters for an inviscid core.

C. The shell filled with silicone oil in the high-frequency range

An improved phase demodulator was used to measure up to 2 MHz, allowing the validity of the core parameters

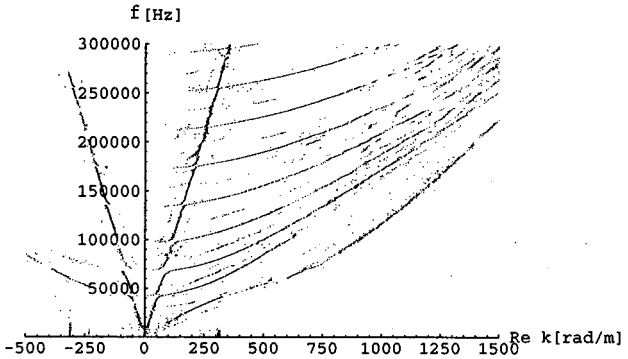


FIG. 8. Projection of the curves presented in Fig. 7 into the real plane.

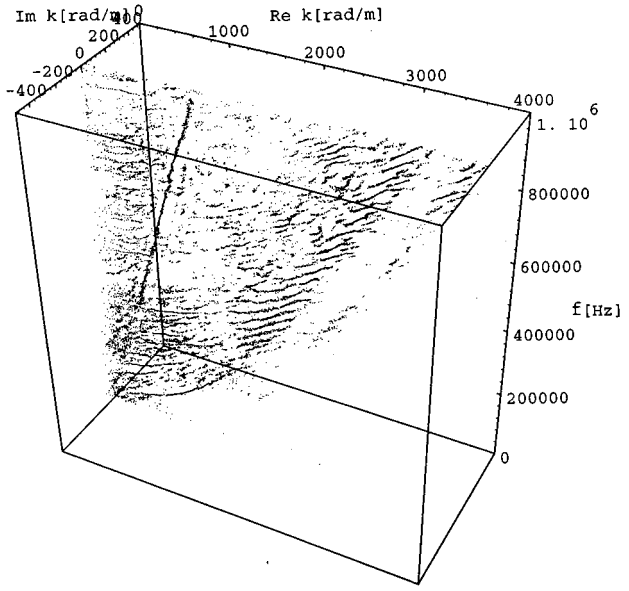


FIG. 9. Measured dispersion curves of approximately the first 25 axisymmetric wave modes in a cylindrical shell filled with silicone oil with a nominal viscosity of 60 000 mm^2/s; temperature: 24.0 °C. Number of samples in space domain: $n_s=120$; $\Delta x_{\text{sample}}=0.75$ mm.

determined below 300 kHz to be verified in a higher-frequency range. Figure 9 shows experimental dispersion curves for the same shell-core configuration (silicone oil with a density of 969.0 kg/m^3 and a nominal viscosity of 60 000 mm^2/s).

Even at an upper frequency which is more than three times the upper frequency of the first series of experiments, the inviscid core model still fits quite well in several portions of the dispersion curves. (See Figs. 9–11). Thus the relation between real wave number and frequency is obviously governed primarily by the velocity c_φ of P waves in the core.

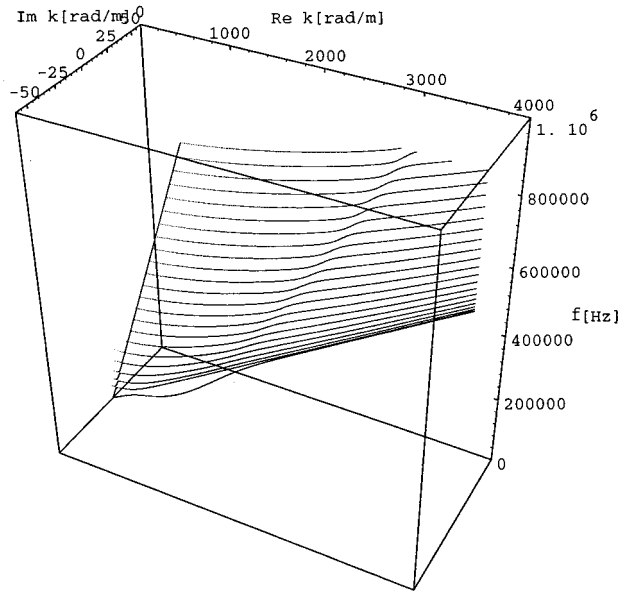


FIG. 10. Theoretical results for a shell filled with an inviscid core material having the same velocity c_φ of P waves as silicone oil.

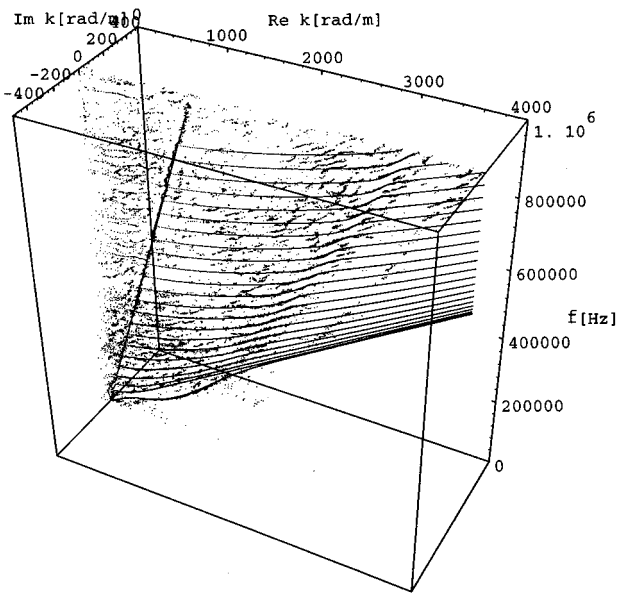


FIG. 11. Experimental and theoretical (inviscid case) results in one diagram.

By viewing the data presented in Fig. 9 from a different angle and magnifying the imaginary axis, one can see the periodically alternating values of the imaginary parts of some dispersion curves. This effect may appear meager; nevertheless, it is clear evidence for the shear elastic behavior of the silicone core, as comparison of Figs. 12 and 13 shows. The region indicated by the arrow in Fig. 13 is connected with the flexural mode of the empty shell (see Part I[16]). The

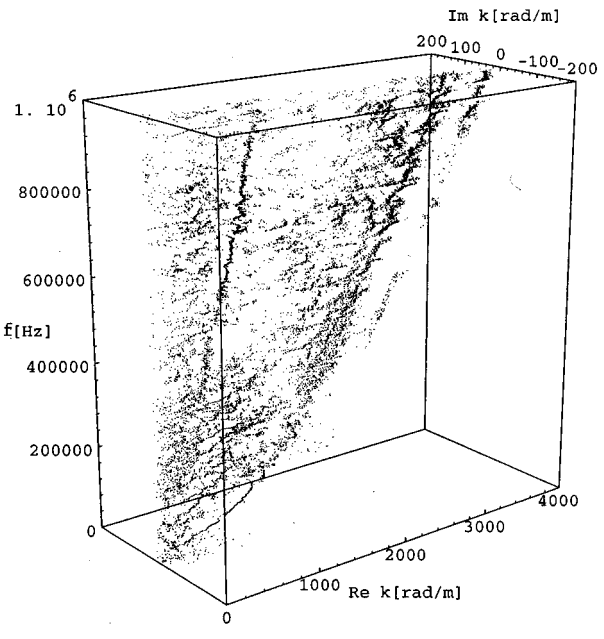


FIG. 12. Measured dispersion curves axisymmetric wave modes in a cylindrical shell filled with silicone oil with a nominal viscosity of 60 000 mm^2/s; temperature: 24.0 °C scaled to emphasize the effect of shear elasticity. Number of samples in space domain: $n_s=120$; $\Delta x_{\text{sample}}=0.75$ mm.

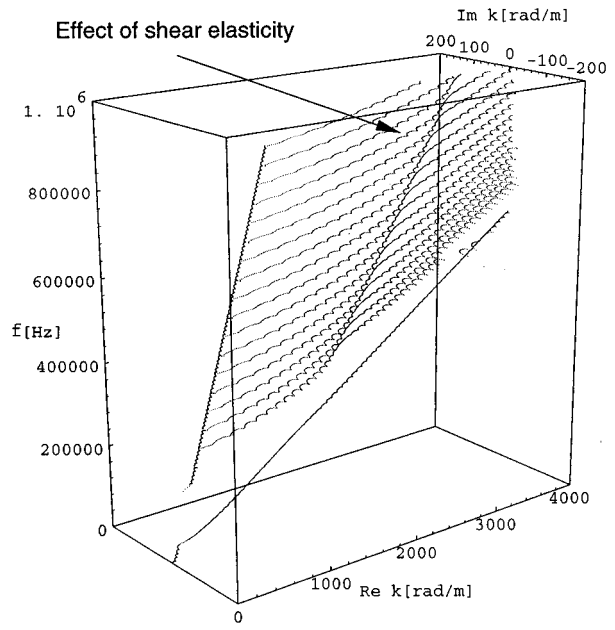


FIG. 13. Theoretical dispersion relation of axisymmetric modes of a shell containing a shear elastic core material having the same velocity c_φ of P waves as silicone oil. (Due to numerical reasons, some lower modes could not be calculated.)

imaginary part of the curves is very small there, a fact which explains the good measurability in that particular range.

D. The shell filled with alcohol (low viscosity)

Dispersion curves are also measured for a shell filled with ethyl alcohol (density: 789 kg/m^3, nominal viscosity: 1.52 mm^2/s at 20 °C). Since the velocity c_φ of P waves in this fluid is lower than in silicone oil, the wavelength is shorter. Therefore more concentric zones of equal pressure are present in the core at one particular frequency, thus increasing the number of detectable modes as shown in Fig. 14.

Again the curves can only be measured clearly in the region of the flexural mode of the empty shell for frequencies above 500 kHz. It is most likely that the complex spectrum estimation method reaches its limits by interpreting the periodically alternating imaginary part of the wave number as noise except in the region mentioned above.

E. The shell filled with polyisobutylene (high viscosity)

Not all measured dispersion diagrams can be explained with the theoretical model introduced in Part I.[16] As an example, the data of a shell filled with polyisobutylene (BASF® Opanol™ b 10) are presented below (Figs. 15–17). Polyisobutylene is a linear, flexible polymer with a glass temperature of -65 °C. Above this temperature the material can be considered an amorphous homogeneous melt. According to BASF®, the density of Opanol™ b 10 equals 890 kg/m^3, the real part of the shear modulus $\mathrm{Re}\,G^*(f=100\ \mathrm{kHz}, t=150\ ^\circ\mathrm{C}) \approx 2\times10^5\ \mathrm{N/m^2}$, and the imaginary part $\mathrm{Im}\,G^*(f=100\ \mathrm{kHz}, t=150\ ^\circ\mathrm{C}) \approx 2\times10^5\ \mathrm{N/m^2}$.

F. A linear elastic aluminum rod

When considering the fluids which have been investigated, it seems that dispersion curves in the complex space have either a very complicated shape which is barely describable by a theoretical model with a limited number of parameters (Figs. 15 and 17), or can be detected only very close to the real plane (Fig. 14). An intrinsic demand of the linear prediction method is that the amplitude of propagating waves needs to be sufficient in order to separate them from surrounding noise. Therefore highly damped modes are difficult to measure.

Linear elasticity can be considered as an extreme case of viscoelasticity having long relaxation times compared with

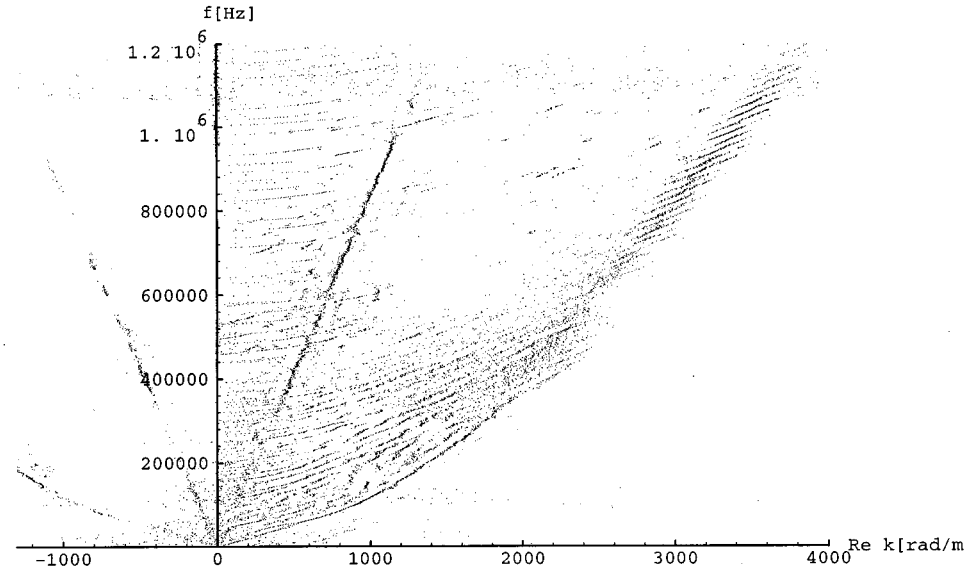


FIG. 14. Measured dispersion curves (projection into the real plane) of axisymmetric wave modes in a cylindrical shell containing ethyl alcohol (density: 789 kg/m^3, nominal viscosity: 1.52 mm^2/s at 20 °C); temperature: 23.8 °C. Number of samples in space domain: $n_s=120$; $\Delta x_{\mathrm{sample}}=0.75$ mm.

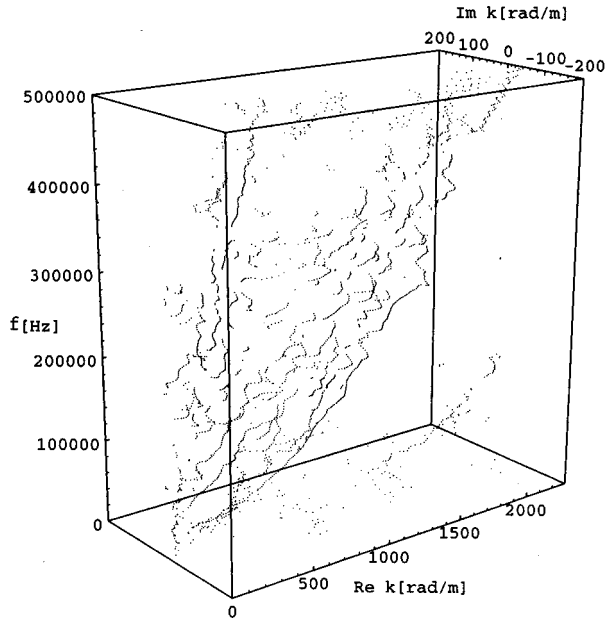


FIG. 15. Dispersion curves of axisymmetric waves in a cylindrical shell filled with polyisobutylene (BASF® b Opanol™ 10), measured at 26.0 °C. Number of samples in space domain: $n_s = 160$; $\Delta x_{\text{sample}} = 0.75$ mm.

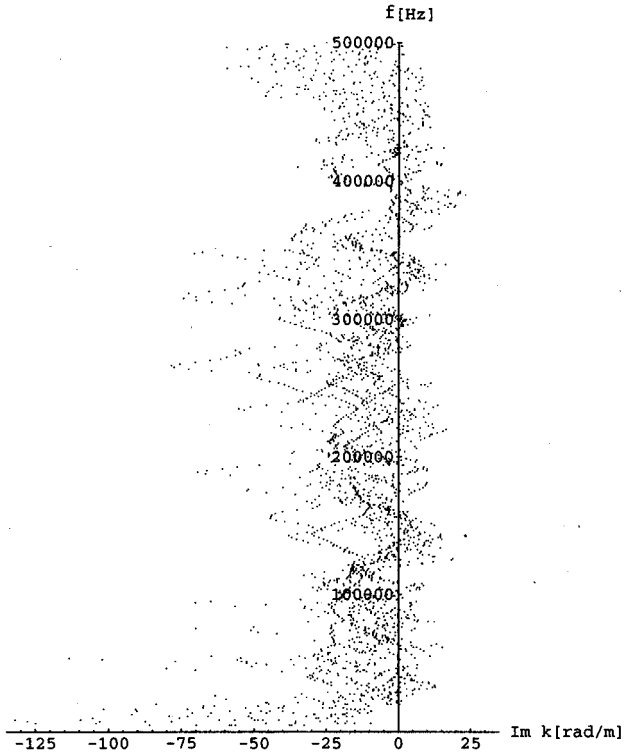


FIG. 17. Projection of the curves presented in Fig. 15 into the imaginary plane.

the deformation period. To ascertain the validity of the theoretical model as well as the computational precision, especially in the range of strong shear elasticity, the dispersion curves of axisymmetric waves in an aluminum rod with circular cross section were also measured (Figs. 18 and 19). This dispersion relation is calculated using the same computer program used for the shell filled with viscoelastic medium (Fig. 20). Therefore the relaxation times for shear and bulk behavior of the core material are set to high values compared with the reciprocal value of the lowest frequency within the range considered. The thickness of the shell wall h is set to a very small value (10 μm). Excellent agreement between material constants found in this manner with values determined by other dynamic methods is found (Fig. 21).

The diameter of the aluminum rod (ALUSWISS® Anticorodal™-112-hard) is 18 mm. With a density of 2751.6 kg/m^3, a Young's modulus of 7.041×10^{10} N/m^2, and a Poisson ratio of 0.35, the best fit with experimental curves is obtained. These values were found by systematically varying one parameter after another until the best fit was achieved.

Since the damping of the waves propagating in an aluminum rod is much lower than in a silicone-filled shell, the dispersion curves of waves reflected at the end of the rod, which have negative wave numbers, are detected as well (see Fig. 18). The fact that energy is propagating in both direc-

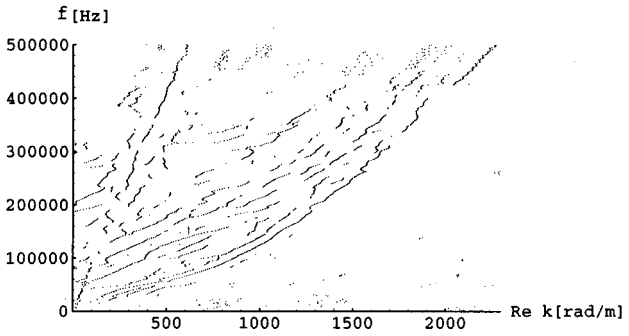


FIG. 16. Projection of the curves presented in Fig. 15 into the real plane.

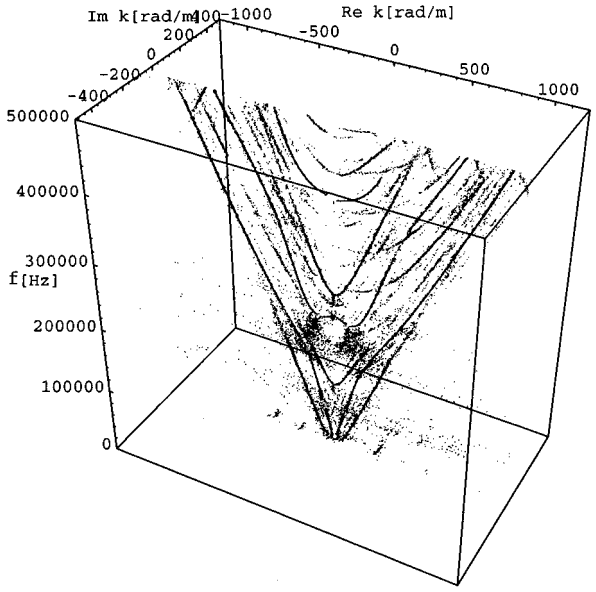


FIG. 18. Measured dispersion relation of axisymmetric wave modes of an aluminum rod with circular cross section (diameter: 18 mm). Number of samples in space domain: $n_s = 130$; $\Delta x_{\text{sample}} = 2.0$ mm.

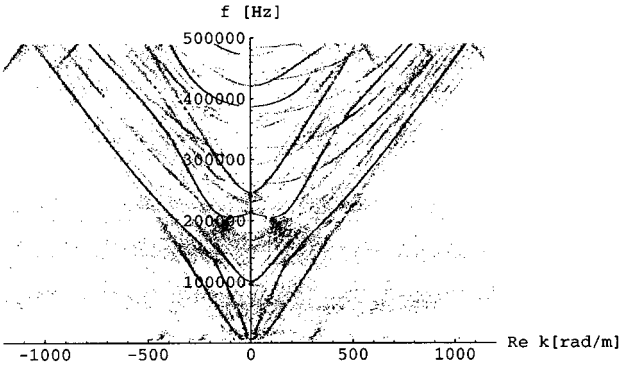


FIG. 19. Projection of the curves presented in Fig. 18 into the real plane.

tions enables the detectability of the phenomenon of ''backward-wave propagation,'' mentioned by Onoe *et al.* in 1962,[24] in which group velocity and phase velocity have opposite signs (Fig. 22). This phenomenon was the subject of extensive discussions in the early 1960s.

The group velocity represents the velocity of the energy carried by a wave. It is proportional to the slope of the dispersion curves in the wave-number-frequency diagram. As Folk[25] explained, the path of a consistent mode cannot change the sign of its slope, but in Fig. 20 the slope of the second lowest mode appears to do so. In fact, this mode ceases to exist beyond the point where its slope reaches zero since no energy can escape from the transmitting piezo transducer. The four branches leaving the real plane at these points represent boundary layers existing only very close to the excitational source. The traces of these curves are also visible in Fig. 22. The short branch with the negative slope (see Figs. 22 and 20) exists due to the energy reflected at the end of the specimen flowing towards the excitation. A first attempt to measure ''backward-wave propagation'' was undertaken by Meitzler[26] in 1965. A detailed investigation of the phenomenon of ''backward wave propagation'' in conjunction with material properties will be published in a later paper.

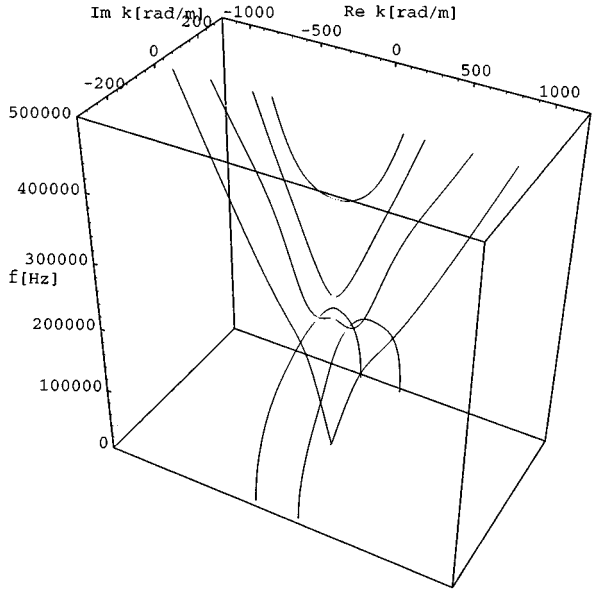


FIG. 20. Theoretical dispersion relation of axisymmetric wave modes of an aluminum rod with circular cross section (diameter: 18 mm). The curves in the imaginary plane are not presented in this diagram.

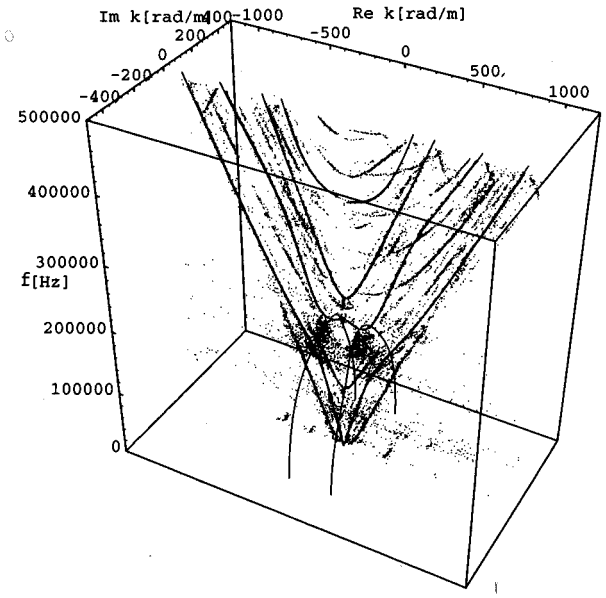


FIG. 21. Experimental and theoretical dispersion curves of an aluminum rod in one diagram.

Apart from axisymmetric waves, the dispersion curves of several nonaxisymmetric modes have also been detected, as seen in Fig. 18. These modes are always present due to imperfections in the excitation, the structure itself, or the support of the specimen.

IV. CONCLUSIONS

A new contactless measurement technique for structural mechanics has been developed by combining modern spec-

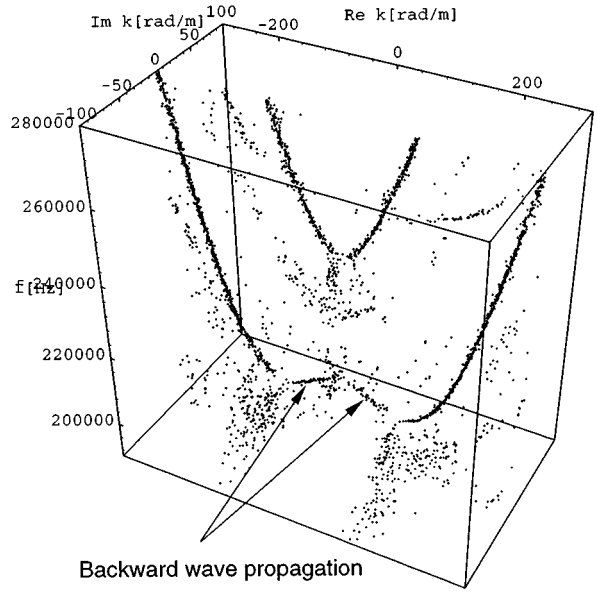


FIG. 22. Measured dispersion relation of axisymmetric wave modes of an aluminum rod with circular cross section, scaled to show the area of ''backward wave propagation.''

trum estimation methods with a high-resolution Laser interferometer. Some aspects of dynamic fluid structure interaction in the high-frequency range, where fluid behavior is not well-known, are investigated. The validity of a multiparameter model describing the dynamic behavior of a thin-walled elastic shell containing various core materials derived in Part I[16] can be verified for several core materials. Since the method is based on propagating waves, the detectability of a mode decreases with increasing damping. A main advantage of the axisymmetry is the fact that the system is mechanically highly decoupled from the surrounding world.

The tools developed in this investigation can be used to explore the topology of the dispersion relation of axisymmetric waves in cylindrical structures. Thus they form a basis for various applications such as the design of fluid sensors or fluid delivery systems, the development of nondestructive material and structure testing methods, the inspection of coatings, and so on. If optimal damping of a structure is required, such as for vibration or noise absorbers, this model can be used to specify a material which fits the purpose best for a specific frequency range. The measurement technique is also used for determining complex material properties of anisotropic materials. These results will be presented in a forthcoming paper. The phenomenon of ''backward wave propagation,'' in which group velocity and phase velocity have opposite signs, is clearly measured. Future work deals with the extension of complex spectrum estimation methods to wave phenomena, propagating in more than one dimension, as well as with the improvement of the signal-to-noise ratio.

ACKNOWLEDGMENTS

The authors would like to express their thanks to Markus Hägeli for the improvement of the Laser interferometer.

[1] C. R. Fuller and F. J. Fahy, ''Characteristics of wave propagation and energy distributions in cylindrical elastic shells filled with fluid,'' J. Sound Vib. **81**, 501–518 (1982).
[2] N. Krause, W. Goldsmith, and J. L. Sackman, ''Transients in tubes containing liquids,'' Int. J. Mech. Sci. **19**, 53–68 (1977).
[3] R. Skalak, ''An extension of the theory of water hammer,'' Paper No. 55-S-18, Hydraulics Division, ASME, Baltimore, MD, 1955.
[4] F. Barez, W. Goldsmith, and J. L. Sackman, ''Longitudinal waves in liquid–filled tubes I (Theory) and II (Experiments),'' Int. J. Mech. Sci. **21**, 213–236 (1979).
[5] L. D. Lafleur and F. D. Shields, ''Low-frequency propagation modes in a liquid-filled elastic tube waveguide,'' J. Acoust. Soc. Am. **97**, 1435–1445 (1995).
[6] J. Zemanek, ''An experimental and theoretical investigation of elastic wave propagation in a cylinder,'' J. Acoust. Soc. Am. **51**, 265–283 (1972).
[7] S. M. Kay, *Modern Spectral Estimation: Theory and Application* (Prentice-Hall, Englewood Cliffs, NJ, 1988).
[8] S. L. Marple, Jr., *Digital Spectral Analysis* (Prentice-Hall, Englewood Cliffs, NJ, 1987).
[9] S. L. Marple, Jr., ''A tutorial overview of modern spectral estimation,'' Proc. IEEE **2152–2157**, 1989.
[10] M. H. Hayes, *Statistical Digital Signal Processing and Modelling* (Wiley, New York, 1996).
[11] S. Oestreich, ''Digitale Spektralanalyse, Theorie und Anwendung bei Vibrationssignalen,'' dissertation, Universität der Bundeswehr München, München, 1992.
[12] M. Uike, T. Uchiyama, and H. Minamitani, ''Comparison of linear prediction methods based on singular value decomposition,'' J. Magn. Reson. **99**, 363–371 (1992).
[13] C. F. Tirendi and J. F. Martin, ''Quantitative analysis of NMR spectra by linear prediction and total least squares,'' J. Magn. Reson. **85**, 162–169 (1989).
[14] G. H. Golub and C. F. Van Loan, ''An analysis of the total least squares problem,'' SIAM (Soc. Ind. Appl. Math.) J. **17**, No. 6 (1980).
[15] G. H. Golub and C. F. Van Loan, *Matrix Computations* (Johns Hopkins U.P., Baltimore, 1989), 2nd ed.
[16] J. Vollmann and J. Dual, ''High-resolution analysis of the complex wave spectrum in a cylindrical shell containing a viscoelastic medium. Part I: Theory and numerical results,'' J. Acoust. Soc. Am. **102**, 896–908 (1997).
[17] J. Dual, M. Hägeli, M. R. Pfaffinger, and J. Vollmann, ''Experimental aspects of quantitative nondestructive evaluation using guided Waves,'' Ultrasonics **34**, 291–295 (1996).
[18] R. Prony, ''Essal Expérimental et Analytique Sur les lois de la Dilatabilité des fluides élastiques et sur celles de la Force expansive de la vapeur de l'eau et de la vapeur de l'alkool, à différentes températures,'' Journal Polytechnique ou Bulletin du Travail fait à L'Ecole Centrale des Travaux Publics, Paris, Premier Cahier, An. III (1795). [Riche, Gaspard-Clair-François-Marie, Baron de Prony, 1755–1839. French mathematician and engineer. Director of cadastral survey (1791); as inspector general of bridges and roads (1805–39), responsible for all civil engineering projects, such as harbor improvements, drainage of Pontine marshes, and straightening course of Po River; invented a friction brake (1821).]
[19] F. B. Hildebrand, *Introduction to Numerical Analysis* (McGraw-Hill, New York, 1956).
[20] S. L. Marple, Jr., ''A New Autoregressive Spectrum Analysis Algorithm,'' presented at the 12th Annual Asilomar Conference on Circuits and Systems, Asilomar, CA, 1978 (unpublished).
[21] D. W. Tufts and R. Kumaresan, ''Estimation of frequencies of multiple sinusoids: Making linear prediction perform like maximum likelihood,'' Proc. IEEE **70**, 975–989 (1982).
[22] S. Van Huffel and J. Vandewalle, *The Total Least Squares Problem, Computational Aspects and Analysis* (Society for Industrial and Applied Mathematics, Philadelphia, 1991).
[23] G. R. Cowper, ''The shear coefficient in Timoshenko's beam theory,'' J. Appl. Mech. **33**, 335–340 (1966).
[24] M. Onoe, H. D. McNiven, and R. D. Mindlin, ''Dispersion of axially symmetric waves in elastic rods,'' J. Appl. Mech. **29**, 729–734 (1962).
[25] R. T. Folk, ''Time dependent boundary value problems in elasticity,'' Ph.D. dissertation, Lehigh University, Bethlehem, Pennsylvania, 1958.
[26] A. H. Meitzler, ''Backward-wave transmission of stress pulses in elastic cylinders and plates,'' J. Acoust. Soc. Am. **38**, 835–842 (1965).

On the measurement of residual stress through changes in critical angle

John J. Ditri[a)]
1401 Disston Street, Philadelphia, Pennsylvania 19111

(Received 10 February 1997; accepted for publication 14 April 1997)

The critical angle method of measuring residual stresses in materials, sometimes referred to as *critical angle reflectivity*, has been reexamined with an emphasis on improving the sensitivity of the technique by a judicious choice of reference material. The sensitivity of the critical angle to prestress level was found to be dependent upon the ratio of phase velocities of the incident wave to that of the critically refracted wave in an unstressed medium. The sensitivity is shown to increase monotonically as this ratio approaches unity. Some numerical results are tabulated for the expected change in critical angle per MPa of prestress for various reference/prestressed material combinations for various materials of industrial interest. © *1997 Acoustical Society of America.*
[S0001-4966(97)02508-3]

PACS numbers: 43.35.Zc, 43.35.Cg [HEB]

INTRODUCTION

The presence of residual or service-induced stresses in structures plays a fundamental role in determining the ultimate useful life of the structure. Many techniques have therefore been developed to measure residual and/or applied stresses. A recent book[1] contains a good summary of, and references to, most of the major techniques developed to date.

One very promising *nondestructive* technique is the use of ultrasonic waves. It has been known since at least the time of Cauchy[2] that the presence of stress in a body causes a small change in the phase velocity of virtually all modes of elastic waves propagating through the body. Since these early discoveries, there have been numerous theoretical and experimental contributions to this field,[3–19] which, in analogy with photoelasticity, has been termed acoustoelasticity.

In the usual first-order treatment of the acoustoelastic effect,[3–19] the velocity of an elastic wave in a prestressed body, v, can be expressed in the form,

$$v_{\alpha\beta}=v^0_{\alpha\beta}(1+K^{ij}_{\alpha\beta}\sigma_{ij}), \tag{1}$$

where the subscripts α and β on the velocity indicate the directions of propagation and polarization of the wave, respectively, with $v^0_{\alpha\beta}$ representing the velocity in the initially unstressed medium. The $K^{ij}_{\alpha\beta}$ are the so-called acoustoelastic coefficients, which relate the change in the phase velocity of a wave propagating in the α direction and polarized in the β direction to the applied stress component σ_{ij} alone. In this first-order treatment of the acoustoelastic effect, the effects of different stress components are directly additive; hence, a summation is implied over repeated Latin indices, $i,j \in\{1,2,3\}$.

Numerical values of particular acoustoelastic coefficients (i.e., those relating to specific waves modes and propagation directions) can be found in much of the previously cited literature on acoustoelasticity in, perhaps, a different form than in (1). A few of these coefficients have been measured experimentally and tabulated, for important industrial metals, by Deputat[3] and his co-workers. They considered the propagation of various wave types propagating both along and perpendicular to uniaxially applied stress fields.

The acoustoelastic coefficients corresponding to different wave types, materials, and propagation directions relative to stress directions are found to vary from one another by as much as two orders of magnitude. In general, however, the coefficients are very small quantities, on the order of 0.1×10^{-5} to 10.0×10^{-5} (MPa^{-1}).

The fact that the change in phase velocity is directly related to the applied/residual stress affords a means of determining the stress state by measurement of the (change in) velocity. Because of the smallness of the acoustoelastic coefficients, this requires extremely accurate measurement apparatus with the capability of measuring nanosecond changes in time of flight to resolve stress changes on the order of 1–10 MPa.

One method of overcoming the smallness of these coefficients is to allow the waves to travel over larger distances and hence to gradually acquire larger time shifts. This comes at the cost of averaging the stress over larger portions of the material, and therefore the spatial resolution of the technique is diminished. Velocity measurement is further complicated when pulse shape changes take place, making the velocity somewhat harder to unambiguously define and measure.

Another method which has been proposed and used to measure the changes in velocity has been termed *ultrasonic refractometry* or *critical angle reflectivity*. This method consists of carefully measuring the critical angle of incidence of a specific wave type on the interface between a reference material and the material whose stress state is to be measured.

The critical angle is related to the phase velocities of the incident and critically refracted waves by the relation

$$\theta_{\mathrm{cr}}=\arcsin(v_1/v_2), \tag{2}$$

where v_1 represents the phase velocity of the incident wave

[a)]Electronic mail: jditri@voicenet.com

and v_2 represent the phase velocity of the refracted wave in the stressed material in the direction of the interface. Note that the presence of prestress in a medium causes it to become slightly anisotropic, and it is assumed that the originally circular slowness profiles of the refracted wave modes do not develop cusps due to the prestress. Under this assumption, it is the value of phase velocity of the refracted wave when propagating in the direction parallel to the interface which determines the critical angle of the incident wave.

The utility of the critical angle technique is not limited to investigations of *stress-induced* changes in material properties. Any other phenomenon which causes a change in wave speed (or attenuation) can be investigated using the method. Rollins[20] and Becker and Richardson,[21] for example, showed that the angle of critical refraction of a Rayleigh wave on a liquid-immersed specimen could be used to evaluate coating properties and adhesion strength for thinly coated specimens, plastic strain levels in cold-worked specimens, the degree of surface hardness in heat-treated steels as well as the changes in modulus of elasticity accompanying nitriding, hydriding, or stress. Rollins[20,22] also observed that the *magnitude* of the reflection factor dip which occurs near the Rayleigh wave critical angle was very sensitive to changes in physical (such as microstructural) properties of the reflecting medium. Becker and Richardson[23] were able to explain this effect by incorporating attenuation losses in the model of the reflecting medium. Becker[24] also showed that the *phase changes* which occur in the reflected beam near the Rayleigh wave critical angle were even more sensitive to the properties of the reflecting medium than the position of the critical angle.

The measurement of critical angle can be accomplished by monitoring the reflected amplitude of the incident wave as a function of the angle of incidence. As specific critical angles corresponding to particular refracted wave modes are reached, the reflected wave amplitude usually undergoes rapid, sharp fluctuations. Deputat[3] and Rollins,[20,22] for example, describe such a reflectometer and give typical results for the reflection from various specimens immersed in a liquid bath. The reflectometers constructed by these authors are claimed to be capable of measuring the critical angle to within ± 1 angular min (0.017°). An alternate method is to use sending and receiving transducers, mounted to variable angle wedges, which are separated by some fixed distance. The angles of the sending and receiving transducers are then increased until the critically refracted wave of interest is received by the receiver. The refracted wave will, of course, radiate into the receiving wedge at the critical angle. Finally, Rollins[25] introduced a ''right angle reflector'' technique which enables the reflection information to be obtained using a single transducer in a pulse-echo mode as long as the specimen is immersed in a liquid.

Practically, the critical angle method of measuring phase velocity has several major advantages over direct time-based measurements. The length over which the stress field is averaged, for example, can be greatly reduced since only that portion of the structure which is directly insonified by the incident beam plays a role, and this is typically on the order of 12–40 mm. Also, the coupling between the incident beam and the specimen is not as important as it is when making travel time measurements because only the *relative* reflection amplitudes are utilized. While the *actual* reflection amplitude of the incident beam at the critical angles carries information concerning the physical properties of the reflecting medium such as attenuation,[23] only the *locations* of the maxima/minima are required for velocity measurements. The physical characteristics of the transducer(s) used in the critical angle experiments such as the aperture size, the acoustic beam profile, and the pulse shape also play only secondary roles. One should, of course, choose transducers whose lateral dimensions are large enough in comparison to the wavelength being generated so as to define a well-collimated beam. This will restrict the wave number spectrum of the incident beam to a narrow range around the central wave number which coincides with the axis of the transducer and which is used to define the incident angle.

The effect of surface roughness on the location and magnitude of the amplitude dip corresponding to a critically refracted Rayleigh wave was studied by Rollins.[20] He found that the effect was not easy to quantify, and that it seemed to depend upon the type of specimen being insonified. One very practical conclusion was that only the relative magnitude of the surface roughness *compared to the incident wavelength* was important. For low-frequency waves, therefore, minor surface roughness does not cause significant shifts in the critical angle positions. For example, he reports results of experiments on copper where the rms surface roughness was varied (by chemical etching) from 4 to 50 μm, and there was no measurable shift in the location of the minimum at even 45 MHz. The measurement resolution was 1 arc minute. The same range of surface roughness produced a roughly 20 arc-minute shift in the location of the Rayleigh wave critical angle on iron, at the same 45 MHz.

It is most common when using the critical angle measurement technique to convert the measured critical angle back into the velocity of the wave in the stressed medium via Eq. (2). The change in wave speed can then be used to determine the stress in the body by using Eq. (1).

The goal of the present letter is threefold.

(i) An expression is obtained which allows direct calculation of the stress by measurement of the change in critical angle of incidence between stressed and unstressed configurations.

(ii) It is shown how the sensitivity of the critical angle to the stress can be increased by proper choice of incident medium relative to stressed medium.

(iii) Some numerical results are tabulated for the change in critical angle which can be expected for unit change in stress for various material combinations.

I. ANALYSIS

Consider the simple goniometric probes illustrated in Fig. 1(a) and (b). In Fig. 1(a), there is a semicylindrical block of reference material, equipped with an angle adjustable ultrasonic transducer, which is attached to a prestressed structure using an appropriate couplant. The reference medium can also be a liquid provided a proper holder is used for the

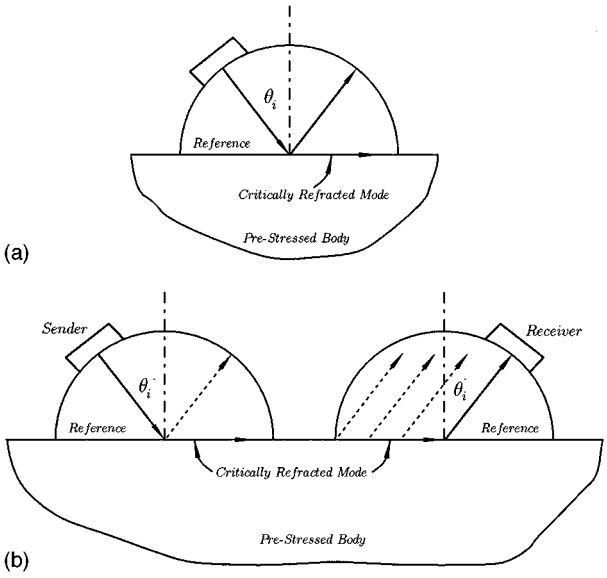


FIG. 1. Two goniometric configurations for measurement of critical angles. (a) Single wedge. (b) Dual wedge.

transducer. The single transducer acts as both the sender and receiver. In Fig. 1(b), there are two ultrasonic transducers mounted to nominally identical variable angle wedges, one for sending the incident wave and the other for monitoring the critically refracted wave.

In either arrangement, a wave is assumed incident in the reference material at an angle θ_i to the common normal of the two media. In the arrangement of Fig. 1(a) the incident wave reflects from the interface and, and after propagating through the wedge, reflects from the free cylindrical surface. The wave then propagates along the same path in the opposite direction until it reaches the sender where its amplitude can be monitored. In the arrangement of Fig. 1(b), the incident wave couples into the wave mode of interest in the specimen material, which then propagates at an angle given by Snell's law. When the incident wave is at the critical angle of incidence, the refracted wave mode propagates along the free surface and, reaching the receiving wedge, radiates into the wedge at the critical angle where it is strongly received by the receiving transducer. In either configuration, the incident wave can be longitudinal or transverse, and its phase velocity is denoted by v_{ref}.

The structure being monitored is assumed to be subjected to homogeneous stresses σ_{ij}. At the critical angle of a particular wave mode β (i.e., longitudinal, transverse, Rayleigh, etc.), the wave in the structure will propagate in the x_1 direction (along the interface) with a speed given by

$$v_{1\beta}=v^0_{1\beta}(1+K^{ij}_{1\beta}\sigma_{ij}), \tag{3}$$

where again summation is implied over the Latin indices, i and j.

From the equation for the critical angle, Eq. (2), the *change* in critical angle due to the applied stresses can be written

$$\Delta\theta_{\mathrm{cr}}=\arcsin\left\{\frac{v_{\mathrm{ref}}}{v^0_{1\beta}(1+K^{ij}_{1\beta}\sigma_{ij})}\right\}-\arcsin\left\{\frac{v_{\mathrm{ref}}}{v^0_{1\beta}}\right\}. \tag{4}$$

The term $K^{ij}_{1\beta}\sigma_{ij}\equiv\Delta v_{1\beta}/v^0_{1\beta}$ actually represents the fractional change in wave speed due to the prestress, and is small even for stress levels near the yield point of most metals.

Under the assumption that $\Delta v_{1\beta}/v^0_{1\beta}\ll 1$, a Taylor series expansion of $\Delta\theta_{\mathrm{cr}}$ about $K^{ij}_{1\beta}\sigma_{ij}=\Delta v_{1\beta}/v^0_{1\beta}=0$ can be obtained from (4) in the form

$$\Delta\theta_{\mathrm{cr}}=-\frac{(v_{\mathrm{ref}}/v^0_{1\beta})}{\sqrt{1-(v_{\mathrm{ref}}/v^0_{1\beta})^2}}K^{ij}_{1\beta}\sigma_{ij}+\left[\frac{(v_{\mathrm{ref}}/v^0_{1\beta})^3}{2\sqrt[3]{1-(v_{\mathrm{ref}}/v^0_{1\beta})^2}}+\frac{(v_{\mathrm{ref}}/v^0_{1\beta})}{\sqrt{1-(v_{\mathrm{ref}}/v^0_{1\beta})^2}}\right](K^{ij}_{1\beta}\sigma_{ij})^2+O[(K^{ij}_{1\beta}\sigma_{ij})^3]. \tag{5}$$

Retaining only the term which is linear in $\Delta v_{1\beta}/v^0_{1\beta}$ in Eq. (5), and defining the function $F(v_{\mathrm{ref}}/v^0_{1\beta})$ as

$$F(v_{\mathrm{ref}}/v^0_{1\beta})\equiv\frac{(v_{\mathrm{ref}}/v^0_{1\beta})}{\sqrt{1-(v_{\mathrm{ref}}/v^0_{1\beta})^2}}, \tag{6}$$

the change in critical angle due to the prestress can be written

$$\Delta\theta_{\mathrm{cr}}=-F(v_{\mathrm{ref}}/v^0_{1\beta})K^{ij}_{1\beta}\sigma_{ij}. \tag{7}$$

Equation (7) is a direct relation between the applied (or residual) stresses and the change in critical angle of an arbitrary wave mode β. It is similar in principle to Eq. (3), relating the change in velocity of a wave to the applied stress with one important exception; namely, the appearance of the functional factor F. This function, which is only dependent upon the ratio $v_{\mathrm{ref}}/v^0_{1\beta}$, serves as a multiplying factor to increase or decrease the sensitivity of the critical angle change to the applied/residual stresses.

Here, $F(v_{\mathrm{ref}}/v^0_{1\beta})$ can be made larger and, hence, the critical angle made more sensitive to the prestress by choosing the reference material of the goniometer such that the phase velocity of the incident mode (v_{ref}) is as close as possible to the phase velocity of mode type β in the *unstressed* state of the material being inspected. Figure 2 is a plot of $F(v_{\mathrm{ref}}/v^0_{1\beta})$ in the range $0\leqslant v_{\mathrm{ref}}/v^0_{1\beta}\leqslant 0.95$. On the top of the scale is the initial critical angle corresponding to the particular value of $v_{\mathrm{ref}}/v^0_{1\beta}$.

Requiring v_{ref} to be close to $v^0_{1\beta}$ is equivalent to requiring that the initial critical angle (i.e., between the reference material and the tested material in its unstressed state) be as large as possible. Size requirements of the emitting and receiving transducers will eventually limit the initial critical angles that can be achieved.

II. UNIAXIAL STRESS STATE

In general, the change in critical angle of any wave mode will depend upon all of the stress components in the stressed medium, as can be seen from Eq. (7). This means that the single critical angle measurement will not supply, in general, sufficient information to obtain the stresses in the specimen.

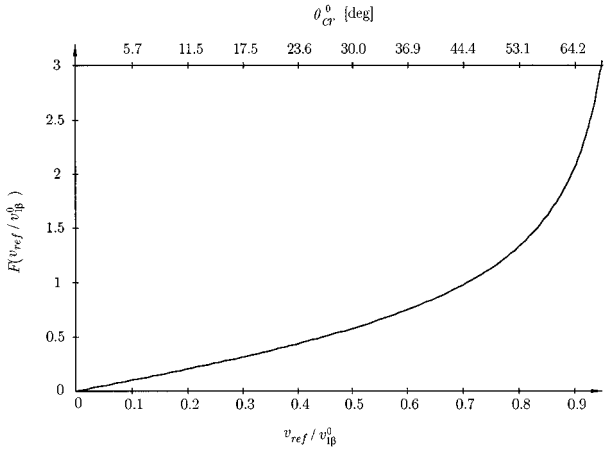


FIG. 2. Plot of function $F(v_{\text{ref}}/v_{1\beta}^0)$ showing the increase in sensitivity possible by proper choice of reference material with respect to the specimen material. The top scale is the initial critical angle (i.e., for the unstressed specimen) associated with the corresponding value of $v_{\text{ref}}/v_{1\beta}^0$.

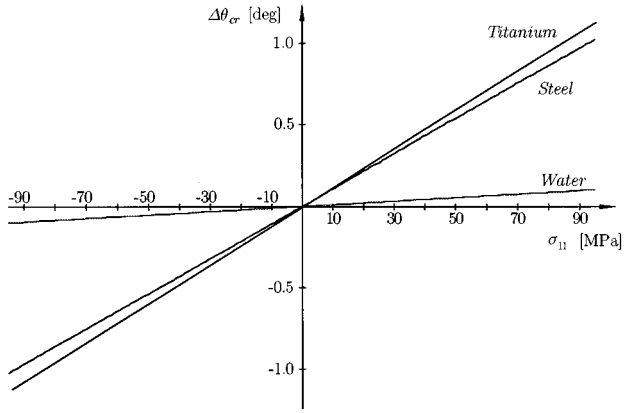


FIG. 3. Change in critical angle of the longitudinal mode between three reference materials and an aluminum specimen versus the uniaxial homogeneous pre-stress level.

In the simplest case of a homogeneous uniaxial prestress in, say, the x_1 coordinate direction (σ_{11}), however, the change in critical angle can be expressed in terms of the single acoustoelastic coefficient, $K_{1\beta}^{11}$, as

$$\Delta\theta_{\text{cr}} = -F(v_{\text{ref}}/v_{1\beta}^0)K_{1\beta}^{11}\sigma_{11}. \tag{8}$$

Measurement of $\Delta\theta_{\text{cr}}$ then suffices to determine σ_{11} from Eq. (8).

Figure 3 is an illustrative example of how the change in critical angle can be made more sensitive to the prestress by proper choice of the reference material. Shown there are plots of the predicted change in critical angle, from Eq. (8), for a longitudinal wave incident from three different reference materials onto a prestressed aluminum half-space. The half-space is assumed to be stressed uniaxially with an initial stress of σ_{11}, and the critical angle is that of the longitudinal wave in the aluminum. At the critical angle, the critically refracted longitudinal wave is assumed to propagate along the interface in the x_1 direction. The range of prestress in the plots is from $-\sigma_Y$ to $+\sigma_Y$, where $\sigma_Y = 95$ MPa is the approximate yield stress of aluminum.

The value of K_{11}^{11} (i.e., $\beta = 1$ by definition for a longitudinal wave traveling in the x_1 direction) was taken from Deputat[3] as $K_{11}^{11} = -7.75\times10^{-5}$ MPa^{-1}. The values used for v_{ref} for the three materials used to generate Fig. 2, as well as others, can be found in Table I.

As can be seen from Fig. 3, the sensitivity of the critical angle to the applied stress can be greatly increased by proper choice of reference medium. The expected change in critical angle for a 1-MPa (145 psi) prestress in aluminum when using titanium as the reference material is 0.01720° (1.032 min), which, compared to a change of only 0.00107°/MPa (0.0644 min/MPa) when using water as the reference medium, represents an over 16-fold increase in sensitivity. Using titanium as a reference material to measure stresses in aluminum means that the initial critical angle is close to 76°.

Tabulated in Table I are the sensitivities of various reference/prestressed material combinations in terms of the expected change in critical angle per MPa of applied stress. In all cases, it is the critical angle of the refracted longitudinal wave due to incidence of a longitudinal wave which is tabulated. The acoustoelastic coefficients of the metals listed in Table I were taken from Deputat,[3] and those of the PMMA were calculated from experimental data presented by Obata *et al.*[19]

As can be seen from the table, choosing the proper reference material can yield critical angle changes on the order of 0.1 to over 1.0 arc-minutes per MPa of prestress in various

TABLE I. Expected change in critical angle, in arc-minutes, per MPa (145 psi) of prestress. The materials listed vertically represent the reference material, and those listed across the table represent stressed materials. The quantities in brackets represent the longitudinal wave phase velocity in [mm/μs] and the acoustoelastic coefficient in [MPa^{-1}], respectively.

Reference material	v_{ref} mm/μs	Aluminum (6.3) (-7.75×10^{-5})	Armco iron (5.9) (-2.69×10^{-5})	Nickel steel 535 (5.7) (-1.13×10^{-5})	Steel M56 (5.7) (-2.06×10^{-5})	Copper (4.7) (-1.88×10^{-5})	PMMA (2.7) (-0.717×10^{-4})
Air	0.33	0.01397	0.00518	0.00225	0.00411	0.00455	0.03035
Water	1.48	0.06439	0.02396	0.01044	0.01904	0.02144	0.16154
Plexiglas	2.7	0.12638	0.04760	0.02089	0.03809	0.04536	
Crown glass	5.3	0.41459	0.18907	0.09815	0.17893		
Nickel	5.6	0.51694	0.27881	0.20465	0.37307		
Steel	5.7	0.56596	0.34606				
Iron	5.9	0.71157					
Titanium	6.1	1.03200					

materials of industrial interest. This type of change should be easily measurable by current goniometric apparatus which is claimed to have accuracy to ± 1 arc-minute.

III. CONCLUSIONS

It has been shown that the sensitivity of the critical angle measurement to prestress can be greatly increased by choosing an incident medium and wave type so that the phase velocity of the incident wave is close to that of the critically refracted wave type when propagating in the unstressed medium. Although only incident and refracted longitudinal waves were discussed, both the incident and refracted wave types can be transverse as well. In addition, it is possible to measure the change in the optimum angle of generation of Rayleigh waves on the prestressed medium and to infer the prestress using the same equations. The longitudinal waves were chosen because they exhibit the largest acoustoelastic coefficients when propagating along the applied stress direction. It may, however, be more difficult to measure the critical angle of the longitudinal waves using the configuration of Fig. 1(a), since, in general, the change in reflection factor is less affected at the longitudinal critical angle than it is at the shear and Rayleigh wave angles. The Rayleigh wave, in particular, gives a very pronounced change in reflection factor when generated. The goniometric configuration of Fig. 1(b) is much more suitable for use with critically refracted longitudinal waves.

[1] *Handbook on Experimental Mechanics*, edited by A. Kobayashi (SEM, New York, 1993).

[2] A. L. Cauchy, ''De la pression ou tension dans un systeme de points materiels,'' Exer. math. **3**, 213 (1828).

[3] J. Deputat, ''Application of the acoustoelastic effect in measurements of residual stresses,'' Arch. Acoust. **15**(1–2), 69–92 (1990).

[4] J. J. Ditri and D. Hongerholt, ''Stress Distribution Determination in Isotropic Materials via Inversion of Ultrasonic Rayleigh Wave Dispersion Data,'' Int. J. Solids Struct. **33**(**17**), 2437–2451 (1996).

[5] M. Hayes and R. S. Rivlin, ''Propagation of a Plane Wave in an Isotropic Elastic Material Subjected to Pure Homogeneous Deformation,'' Arch. Ration. Mech. Anal. **8**(15), 15–22 (1961).

[6] M. Hayes and R. S. Rivlin, ''Surface Waves in Deformed Elastic Materials,'' Arch. Ration. Mech. Anal. **8**, 358–380 (1961).

[7] M. Hirao, H. Fukuoka, and K. Hori, ''Acoustoelastic Effect of Rayleigh Surface Wave in Isotropic Material,'' J. Appl. Mech. **48**, 119–124 (1981).

[8] M. Hirao, H. Fukuoka, and Y. Murakami, ''Resonance Acoustoelasticity Measurement of Stress in Thin Plates,'' Res. Nondestr. Eval. **4** 127–138 (1992).

[9] D. S. Hughes and J. L. Kelley, ''Second-Order Elastic Deformation of Solids,'' Phys. Rev. **92**(5), 1145–1149 (1951).

[10] D. Husson and G. S. Kino, ''A Perturbation Theory for Acoustoelastic Effects,'' J. Appl. Phys. **53**(11), 7250–7258 (1982).

[11] D. Husson, ''A Perturbation theory for the Acoustoelastic Effect of Surface Waves,'' J. Appl. Phys. **57**(5), 1562–1568 (1985).

[12] T. D. Murnaghan, *Finite Deformation of an Elastic Solid* (Wiley, New York, 1951).

[13] R. W. Ogden and D. A. Sotiropoulos, ''On Interfacial Waves in Pre-Stressed Layered Incompressible Elastic Solids,'' Proc. R. Soc. London, Ser. A **450**, 319–341(1995).

[14] Y. H. Pao, W. Sachse, and H. Fukuoka, ''Acoustoelasticity and ultrasonic measurement of residual stress,'' in *Physical Acoustics*, edited by W. P. Mason and R. N. Thurston (Academic, New York, 1984), Vol. 17, pp. 62–143.

[15] D. A. Sotiropoulos and C. G. Sifniotopoulos, ''Interfacial Waves in Prestressed Incompressible Elastic Interlayers,'' J. Mech. Phys. Solids **43**(3), 365–387 (1995).

[16] T. Tokuoka and Y. Iwashimizu, ''Acoustical Birefringence of Ultrasonic Waves in Deformed Isotropic Elastic Media,'' Int. J. Solids Struct. **4**, 383–389 (1968).

[17] R. A. Toupin and B. Bernstein, ''Sound waves in deformed perfectly elastic materials. Acoustoelastic effect,'' J. Acoust. Soc. Am. **33**, 216–225 (1961).

[18] A. Tverdokhlebov, ''On the acoustoelastic effect,'' J. Acoust. Soc. Am. **73**, 2006–2012 (1983).

[19] M. Obata, H. Shimada, and T. Mihara, ''Stress Dependence of Leaky Surface Wave on PMMA by Line-focus-beam Acoustic Microscope,'' Exp. Mech. 34–39 (1990).

[20] F. R. Rollins, Jr., ''Critical Angle Reflectivity—A Neglected Tool for Material Evaluation,'' Mater. Eval. **24**, 683–689 (1966).

[21] F. L. Becker and R. L. Richardson, in *Research Techniques in Nondestructive Testing*, edited by R. S. Sharpe (Academic, New York, 1970), Vol. 1, Chap. 4, p. 91.

[22] F. R. Rollins, Jr., ''Ultrasonic Reflectivity at a Liquid-Solid Interface near the Angle of Incidence for Total Reflection (Method of measuring degree of cold work, preferred orientation and elastic moduli),'' Appl. Phys. Lett. **7**(8), 212–214 (1965).

[23] F. L. Becker and R. L. Richardson, ''Critical angle reflectivity,'' J. Acoust. Soc. Am. **45**, 793–794 (1969).

[24] F. L. Becker, ''Phase Measurements of Reflected Ultrasonic Waves near the Rayleigh Critical Angle,'' J. Appl. Phys. **45**(1), 199–202 (1971).

[25] F. R. Rollins, Jr., ''Ultrasonic examination of liquid–solid boundaries using a right angle reflector technique,'' J. Acoust. Soc. Am. **44**, 431–434 (1968).

Scattering of elastic waves from thin shapes in three dimensions using the composite boundary integral equation formulation

Yijun Liu
Department of Mechanical, Industrial and Nuclear Engineering, University of Cincinnati, P.O. Box 210072, Cincinnati, Ohio 45221-0072

Frank J. Rizzo
Department of Aerospace Engineering and Engineering Mechanics, Iowa State University, Ames, Iowa 50011

(Received 18 December 1996; accepted for publication 28 April 1997)

In this paper, the composite boundary integral equation (BIE) formulation is applied to scattering of elastic waves from thin shapes with small but *finite* thickness (open cracks or thin voids, thin inclusions, thin-layer interfaces, etc.), which are modeled with *two surfaces*. This composite BIE formulation, which is an extension of the Burton and Miller's formulation for acoustic waves, uses a linear combination of the conventional BIE and the hypersingular BIE. For thin shapes, the conventional BIE, as well as the hypersingular BIE, will degenerate (or nearly degenerate) if they are applied *individually* on the two surfaces. The composite BIE formulation, however, will not degenerate for such problems, as demonstrated in this paper. Nearly singular and hypersingular integrals, which arise in problems involving thin shapes modeled with two surfaces, are transformed into sums of weakly singular integrals and nonsingular line integrals. Thus, no finer mesh is needed to compute these nearly singular integrals. Numerical examples of elastic waves scattered from penny-shaped cracks with varying openings are presented to demonstrate the effectiveness of the composite BIE formulation. [S0001-4966(97)05008-X]

PACS numbers: 43.35.Zc [HEB]

INTRODUCTION

The modeling of thin shapes or thin bodies (including shell-like structures, thin inclusions, thin voids or open cracks in solids, thin-layer interfaces in composites, etc.) is of increasing importance and interest in the fields of acoustics, elastodynamics, and fluid-structure interactions. It is well known that the conventional boundary integral equation (CBIE) formulations, as well as the hypersingular BIE (HBIE) formulations for acoustic and elastic wave problems, will degenerate or break down when they are applied to thin bodies.[1–3] This degeneracy or breakdown is due to the fact that the equation on one side of the thin shape is almost the same as the equation on the other side, see Fig. 1. Eventually the two equations from the two sides become identical when the thickness of the thin shape approaches zero. In the literature, the BIE formulations developed for dealing with the thin shape breakdown can be divided into two groups: one applied to the *one surface model* and one to the *two surface model* of the thin shape. One exception to this classification is the multidomain method[4,5] which can be used for both one surface and two surface models.

The one surface model of a thin shape is an idealized model in which the middle surface of the thin shape is chosen and usually the hypersingular BIE is applied in terms of the pressure jump (for acoustic problems) or the displacement jump (for elastodynamic problems) across the thin shape. This approach has been applied successfully to the problems of scattering and radiation of *acoustic* waves from thin rigid bodies, see, e.g., Refs. 6–11. The scattering of *elastic* waves from planar cracks in three-dimensional elastic medium is also studied using this one surface approach.[7,12,13] This single surface model for thin shapes is efficient in modeling and computation, as long as the hypersingular integrals in the BIE formulations are dealt with properly. However, there are some drawbacks and limitations with this approach. For example, effects of the varying thickness of a thin body or the opening of a crack cannot be studied using this simple model. The use of the jump terms (e.g., pressure jump), instead of the usual boundary variables (pressure), in the hypersingular BIE may also present some inconveniences in the study of a regular body with thin shapes, where the regular BIE and the hypersingular BIE for thin shapes need to be used together.[10]

The two surface model of a thin shape is a more realistic model in that the geometry of the thin shape is not altered. The effects of the thinness and other details of the thin shape can be addressed easily using the two surface model. Early studies of scattering of *acoustic* waves from thin shapes (rigid disks and fluid inclusions with varying thickness in acoustic media) can be found in Refs. 2, 3, and 14. In these studies, the conventional BIE is used on one surface of the thin shape and the hypersingular BIE on the other surface to obtain a nondegenerate system of equations, see Fig. 2. The use of the two surface model is more computation intensive than the one surface model, but it does provide more information about the physics of the problem than the one surface model, as demonstrated in these studies. For example, in the case of acoustic scattering from a rigid disk, the scattered fields from the disk of a finite thickness $2h$ will depart from those fields from the disk of zero thickness, when h is greater than 5% of the radius of the disk. This difference is more

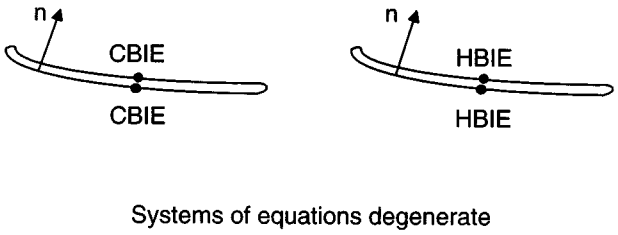


FIG. 1. Degeneracy of the CBIE and HBIE for thin shapes.

pronounced in the near field or when the incident wave is in the direction parallel to the surfaces of the disk, especially if the disk is a ''rippled'' one.[2,3] All these studies on acoustic waves from thin shapes show that the two surface model using the combination of CBIE and HBIE can provide a general BIE formulation and useful information in the analysis of thin shapes. The same BIE formulation can be applied to thin shapes with zero thickness or small thickness, as well as to bulky bodies.

The multidomain method[4,5] is simple and straightforward. It can be applied to thin shapes with zero thickness or nonzero thickness. In this method, the difficulty of dealing with the degeneracy of the conventional BIE for thin shapes is avoided by introducing an imaginary interface to divide the domain into an interior subdomain and an exterior subdomain. The conventional BIE is applied in the two subdomains with meshes on the thin body and interface surfaces, and the continuity conditions are imposed on the interface. This results in a larger system of equations and hence increases the burden of computation, because of the introduction of the imaginary interface. The multidomain method has been used effectively for problems of radiation and scattering of *acoustic* waves from thin rigid bodies.[4,5] Although this approach may be applied to thin body problems posed in a more general setting (e.g., with varying thickness), ''it may not be an ideal tool for the thin-body radiation and scattering problem in which a relatively large imaginary interface surface is usually required.'' [9]

In this paper, the idea of using a combination of the conventional BIE and the hypersingular BIE for thin shapes[2,3,14] is extended to the problems of scattering and radiation of *elastic waves* from thin shapes in elastic media. An alternative form of the CBIE and HBIE combination is used here for the thin shapes, that is, a linear combination of the CBIE and HBIE applied on both surfaces of the thin shapes, see Fig. 2. This composite BIE formulation for elastic wave problems is an extension of the well-known Burton and Miller's BIE formulation[15] for acoustic wave problems, which can remove the fictitious eigenfrequency difficulties in exterior problems. It has also been shown analytically,[3] and will be demonstrated numerically in this paper, that this composite BIE formulation will not degenerate when applied to thin shapes modeled with two distinct surfaces. Thus Burton and Miller's type of BIE formulation can overcome the fictitious eigenfrequency difficulty and the thin body break down difficulty at the same time, being likely the most sound and robust BIE formulations for the acoustic and elastic wave problems.

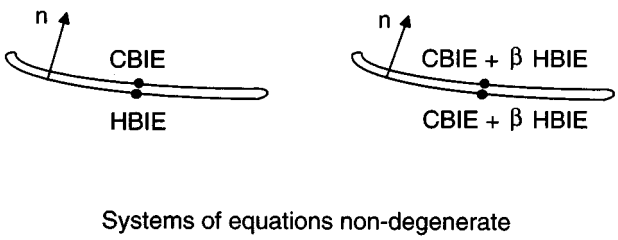


FIG. 2. Nondegeneracy of the combinations of the CBIE and HBIE for thin shapes.

For elastic wave problems, the conventional BIE formulation, which contains strongly singular kernels, has been applied successfully to *bulky-shaped* voids or inclusions for almost a decade (see, e.g., Refs. 16–19). On the other hand, the hypersingular BIE, which contains hypersingular kernels and usually written on one surface of a crack and in terms of the crack opening displacement (COD), has been employed almost exclusively to the problem of scattering from *closed* or *tight* cracks (see, e.g., Refs. 7, 12, and 13). To the authors' best knowledge, no BIE solutions have been reported in the literature for the problem of *elastic* wave scattering from *open* cracks or *thin* voids, or *thin* inclusions. On the other hand, many real problems and experimental calibrations deal with open cracks, notches, and rough cracks with asperities, for which the ideal, *one surface model* is insufficient. Further, many situations of interest involve scattering from thin inclusions and thin-layer interfaces where a shell-like one surface model or lumped-parameter model of the thin region is inadequate. The present study aims to fill this gap and provide a BIE modeling tool for thin shapes with more realistic geometry, by using the composite BIE formulation.

This composite BIE formulation applied in this study was originally developed in Ref. 20 to overcome the fictitious eigenfrequency difficulty existing in the conventional BIE formulation of the *exterior elastic wave* problems. To avoid the hypersingular integrals, the hypersingular BIE is recast in a form in which all integrals are at most weakly singular and thus no special numerical schemes are needed. Nearly singular and nearly hypersingular integrals in this composite BIE formulation, which arise when parts of the boundary surface become close to one another, as is the case for thin shapes, are transformed into sums of weakly singular integrals and nonsingular line integrals. Thus, no finer mesh is needed to deal with these nearly singular integrals. In order to demonstrate the effectiveness of the composite BIE formulation for problems with thin shapes, numerical examples of scattering from penny-shaped cracks with varying openings are given for both longitudinal and transverse incident waves. Results from these numerical example show that the composite BIE formulation is very stable for all ranges of the thinness of a thin shape, even when the two surfaces touch each other. It is also shown, as already demonstrated in acoustics,[2,3,14] that scattered fields for an open crack with an opening of $2h$ will depart noticeably from those fields for a tight crack when h is larger than 5% of the radius of the crack, especially for plane shear waves with an oblique incident angle. All these suggest that the composite BIE formulation, as proposed in this paper for analyzing thin shapes, is

not only robust but also useful in providing information about the physics of such problems.

I. THE COMPOSITE BIE FORMULATION

Consider a region (a body or a void) with the boundary S and immersed in a 3-D infinite, linear elastic medium. The conventional BIE (CBIE) for problems of scattering of time-harmonic waves in the *exterior* domain can be written in the following weakly singular form[18] (index notation is used in this paper):

$$u_i(P_0)+\int_S [T_{ij}(P,P_0)-\overline{T_{ij}}(P,P_0)]u_j(P)dS(P) + \int_S \overline{T_{ij}}(P,P_0)[u_j(P)-u_j(P_0)]dS(P) = \int_S U_{ij}(P,P_0)t_j(P)dS(P)+u_i^I(P_0), \quad \forall P_0 \in S, \quad (1)$$

where u_i and t_i are the total displacement and traction vectors, respectively, U_{ij} and T_{ij} the two dynamic kernels (dependence on the frequency is implied), $\overline{T_{ij}}$ the static kernel, u_i^I the displacement vector of the incident wave, P and P_0 the field and source points, respectively. For interior problems, the free terms $u_i(P_0)$ and $u_i^I(P_0)$ in Eq. (1) will not be present.

The hypersingular BIE (HBIE), or traction BIE, can be written in the following weakly singular form,[20]

$$t_i(P_0)+\int_S \overline{H_{ij}}(P,P_0) \times\left[u_j(P)-u_j(P_0)-\frac{\partial u_j}{\partial \xi_\alpha}(P_0)(\xi_\alpha-\xi_{0\alpha})\right]dS(P) + \int_S [H_{ij}(P,P_0)-\overline{H_{ij}}(P,P_0)]u_j(P)dS(P) + E_{jkpq}e_{\alpha q}\frac{\partial u_p}{\partial \xi_\alpha}(P_0)\int_S [\overline{K_{ij}}(P,P_0)n_k(P) + \overline{T_{ji}}(P,P_0)n_k(P_0)]dS(P) = \int_S [K_{ij}(P,P_0)+\overline{T_{ji}}(P,P_0)]t_j(P)dS(P) - \int_S \overline{T_{ji}}(P,P_0)[t_j(P)-t_j(P_0)]dS(P)+t_i^I(P_0), \quad \forall P_0 \in S, \quad (2)$$

in which H_{ij} and $\overline{H_{ij}}$ are the dynamic and static hypersingular kernels, respectively, K_{ij} and $\overline{K_{ij}}$ another pair of singular kernels, n_k the components of the normal, E_{ijkl} the elastic modulus tensor, ξ_α and $\xi_{0\alpha}$ ($\alpha=1,2$) the two (tangential) coordinates of the points P and P_0, respectively, in a local curvilinear coordinate system defined on the surface S and $e_{\alpha k}=\partial\xi_\alpha/\partial x_k$ evaluated at $\xi_\alpha=\xi_{0\alpha}$ ($k=1,2,3$). Details of the derivation, notation and expressions of all the kernel functions in Eq. (2) can be found in Ref. 20. This HBIE was originally proposed to deal with problems of scattering and radiation from *bulky-shaped* objects.

Both the CBIE and HBIE will degenerate, i.e., become unsolvable or ill conditioned, when they are applied *individually* to thin shapes, e.g., to crack-like problems (imagine that a bulky void becomes a thin void or open crack) as well as true-crack problems,[2,3] see Fig. 1. This degeneracy is manifested by the fact that algebraic equations generated from the BIE on one surface of the crack are (almost) the same as the equations generated on the other surface of the crack. The condition number of the system of equations will increase sharply as the two surfaces of the crack become close. As discussed in Refs. 2 and 3, one remedy to this degeneracy associated with the crack-like (or thin-body) problem is to apply CBIE on one surface of the crack and HBIE on the other surface. This approach, using two surfaces in the model, has been demonstrated to be very effective for *acoustic* wave scattering from thin rigid screens[2] and thin inclusions.[3] Alternatively, and perhaps more advantageous due to the symmetry, it was found that Burton and Miller's composite BIE formulation,[15] using a linear combination of the CBIE and the HBIE as shown symbolically by

$$\text{CBIE}+\beta\text{HBIE}=0 \quad (3)$$

(β is the coupling parameter) will not degenerate when it is applied on both surfaces of a thin void or thin body,[3] see Fig. 2. This composite formulation was originally proposed in Ref. 15 to overcome the fictitious eigenfrequency difficulty (FED) existing in the BIE formulations for exterior *acoustic* wave problems. Recent implementations of this composite BIE formulation to deal with the FED, with weakly singular forms of the HBIEs as key ingredients, can be found in Ref. 20 for *elastic* wave problems and in Ref. 21 for *acoustic* wave problems.

The composite BIE formulation (3), in the context of *elastic* wave problems, using the linear combination of Eq. (1) (CBIE) and Eq. (2) (HBIE), is employed in this study to investigate the problem of elastic wave scattering from thin shapes (e.g., open cracks) in 3 D. This composite BIE formulation is quite general and can be applied to many other thin body problems, such as thin inclusions, thin-layer interfaces in composites and so on. The main objective here is to demonstrate the advantages of this composite formulation which can overcome both the fictitious eigenfrequency difficulty and the thin body difficulty in the conventional BIE formulation. The acoustic wave counterpart of this formulation (Burton and Miller's BIE formulation) has the same feature. Thus this composite BIE formulation can provide unique solutions for scattering from bodies with any shapes (including thin shapes) and at any frequencies. This means that one does not need to switch BIE formulations when dealing with fictitious eigenfrequency and thin body problems.

II. NEARLY SINGULAR AND HYPERSINGULAR INTEGRALS

To apply the composite BIE formulation to thin-body problems, one has to overcome another difficulty, i.e., the *nearly singular* and *nearly hypersingular integrals* which

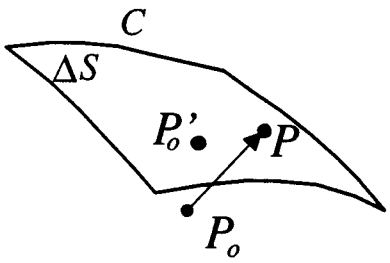


FIG. 3. Nearly singular integrals on surface ΔS enclosed by line C.

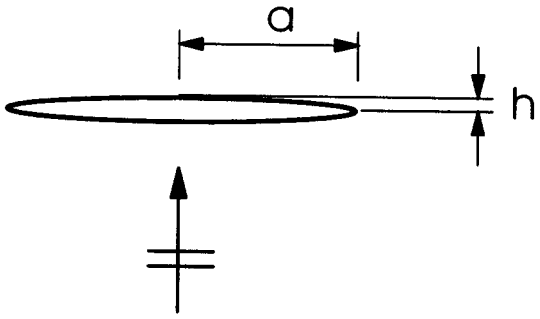


FIG. 4. A penny-shaped open crack with normal incident longitudinal wave.

arise when collocation point is on one surface of the thin body and integration need to be performed on the other nearby surface. Many numerical schemes for computing the nearly singular integrals can be found in the literature of the boundary element method (BEM). Among all the available methods, the line integral approach, i.e., transforming the nearly singular integral into a sum of weakly singular integrals and nonsingular line integrals, is believed to be the most effective and efficient one.[22] It is the method adopted and implemented in this study.

A typical nearly singular integral in Eq. (1) (CBIE) is the one with the stress kernel function T_{ij} and integrated on a surface ΔS with source point P_0 nearby, Fig. 3. Here ΔS can be one element or several elements on the surface S. This nearly singular integral can be dealt with by adding and subtracting terms in the following manner:

$$\begin{aligned}
&\int_{\Delta S} T_{ij}(P,P_0)u_j(P)dS(P) \\
&\quad = \int_{\Delta S} [T_{ij}(P,P_0) - \overline{T_{ij}}(P,P_0)]u_j(P)dS(P) \\
&\qquad + \int_{\Delta S} \overline{T_{ij}}(P,P_0)u_j(P)dS(P) \\
&\quad = \int_{\Delta S} [T_{ij}(P,P_0) - \overline{T_{ij}}(P,P_0)]u_j(P)dS(P) \\
&\qquad + \int_{\Delta S} \overline{T_{ij}}(P,P_0)[u_j(P) - u_j(P_0')]dS(P) \\
&\qquad + u_j(P_0') \int_{\Delta S} \overline{T_{ij}}(P,P_0)dS(P), \qquad (4)
\end{aligned}$$

in which P_0' is the closest point on ΔS to P_0 (an image point of P_0 on ΔS), see Fig. 3. The first two integrals in (4) are now at most nearly weakly singular and can be computed using the normal quadrature. The last integral in (4) can be transformed into line integrals as follows.[22]

$$\begin{aligned}
\int_{\Delta S} \overline{T_{ij}}(P,P_0)dS(P) = I_\Omega(P_0)\delta_{ij} + \frac{1}{4\pi}\epsilon_{ijk}\oint_C \frac{1}{r}dx_k \\
+ \frac{1}{8\pi(1-\nu)}\epsilon_{jkl}\oint_C r_{,ik}\,dx_l, \qquad (5)
\end{aligned}$$

where $r = |\overrightarrow{P_0P}|$, ν is Poisson's ratio, C is the boundary curve of ΔS, ϵ_{ijk} is the permutation tensor, and

$$I_\Omega(P_0) = -\frac{1}{4\pi}\int_{\Delta S}\frac{1}{r^2}\frac{\partial r}{\partial n}\,dS$$

is a solid angle integral which can also be evaluated using a line integral. All these line integrals are nonsingular at all since the source point P_0 is always off the contour C. The nearly *hypersingular* integrals presented in Eq. (2) (HBIE) can be dealt with in a similar way. The expressions of the line integrals for integrals involving the *static* hypersingular kernels can be found in Ref. 22.

Using this line integral approach to deal with the nearly singular integrals is very efficient in computation. One does not need to use more elements in the model in order to handle these nearly singular integrals. It was found that the CPU time used to compute these nearly singular integrals using the line integral approach is only a fraction of that when using many subdivisions on the original surface elements, in achieving the same accuracy.[22]

III. NUMERICAL EXAMPLES

To demonstrate the effectiveness of the proposed composite BIE formulation for problems involving thin shapes, the problem of elastic wave scattering from *open cracks* is studied, for which an analytical solution[23] is available when the opening is zero.

In the first case, a penny-shaped open crack with radius a and thickness $2h$ in a 3-D elastic medium is impinged upon by a plane *longitudinal wave* in the normal direction, see Fig. 4. The scattering cross section for various openings at the nondimensional (shear) wave numbers $K_T a = 1,2,\ldots,6$ are computed using the composite BIE and compared with the analytical solution[23] (only a limited number of data points are available) which is valid for true tight cracks (with zero opening).

Figure 5 shows the results for a very small opening ($h = 0.000001a$) using an increasing number (M) of *nonconforming* quadratic boundary elements[7,20,21] on the two surfaces of the crack. As expected, the BIE solutions are converging to the analytical solution for the tight crack. The small difference is probably due to the fact that the singularity feature of the field near the crack tip is not built in the boundary elements in that region. The singularity feature near the tip of an *open* crack, which depends on the ''open-

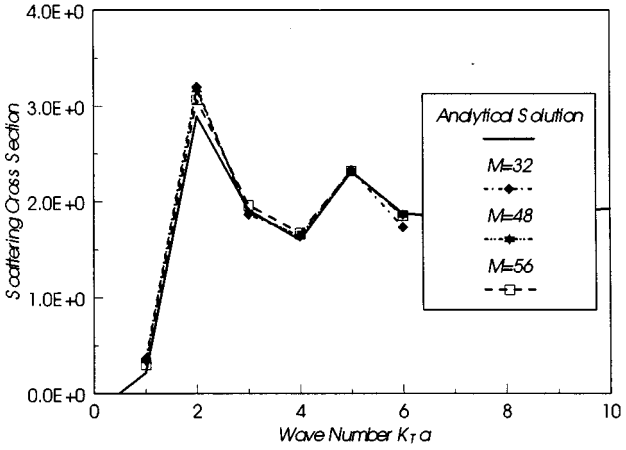


FIG. 5. Convergence of BIE solutions for $h=0.000001a$, longitudinal wave.

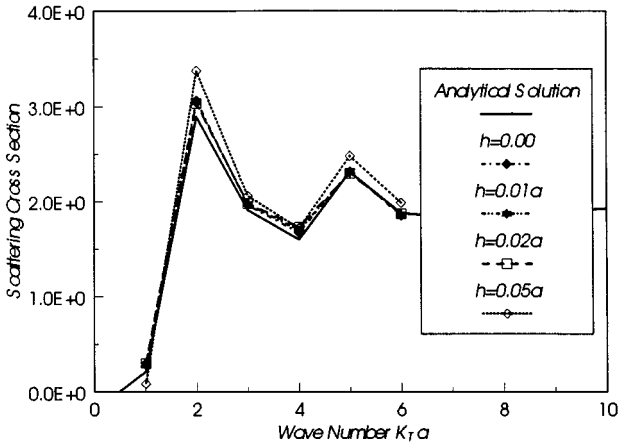


FIG. 7. BIE solutions using 56 elements at different openings, longitudinal wave.

ing angle'' at the crack tip, is difficult to implement and hence is not attempted in this study for elements near the crack tips.

Figure 6 is a similar plot, but for a larger opening ($h=0.02a$). The BIE results are still converging to the tight crack solution. This shows that the small opening does not have a noticeable effect on the far-field data, at least for a plane longitudinal wave in the normal direction of the crack.

Figure 7 shows the results for the crack with four different openings using 56 boundary elements. Noticeable departure of the BIE solution from the analytical solution for tight cracks is observed at the opening $h=0.05a$. This is the departure point for the longitudinal wave in the normal direction. When $h>0.1a$, further departures from the tight crack solution can be observed (not shown here), and one can choose either CBIE *alone* or HBIE *alone* to solve the problem since the degeneracy associated with them is not so severe.

In the next case, the open penny-shaped crack is impinged upon by plane shear waves at different angles of incidence, see Fig. 8. The scattering cross sections are computed, using the composite BIE, at each angle (ranging from 0° to 90° with a 15° increment) of incidence and compared with the analytical solution given in Ref. 23 for a tight crack.

Figure 9 shows the convergence of the composite BIE solutions for the opening $h=0.000001a$ and at the wave number $K_T a=4$. The small gap between the BIE solutions and the analytical solution maybe once again due to the singularity near the crack tip which is not implemented in the boundary elements.

Figure 10 is a plot of the BIE solutions using 56 boundary elements for different openings and at $K_T a=4$. Unlike the case of normal incidence of longitudinal wave (Fig. 7), significant departure of the BIE solution at $h=0.05a$ from the analytical solution for tight cracks is observed. Similar phenomenon is present at a lower wave number ($K_T a=3$), as shown in Fig. 11.

In all the computations, the systems of equations using the composite BIE formulation are well behaved (condition numbers are low). The choice of the coupling parameter β used in Eq. (3) is not so restrictive and values between -1 to $+1$ are found to be adequate.

IV. CONCLUSION

The composite boundary integral equation formulation, using a linear combination of CBIE and HBIE, is proposed for elastic wave problems involving thin shapes (open cracks or thin voids, thin inclusions, thin layer interfaces, etc.) modeled with *two surfaces*. The BIE formulation is very stable no matter how close the two surfaces are, and no undue

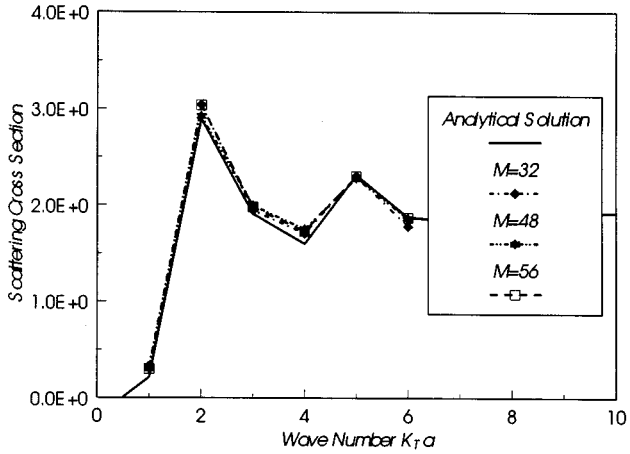


FIG. 6. Convergence of BIE solutions for $h=0.02a$, longitudinal wave.

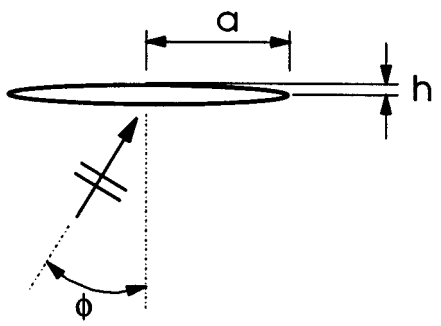


FIG. 8. A penny-shaped open crack with oblique incident shear waves.

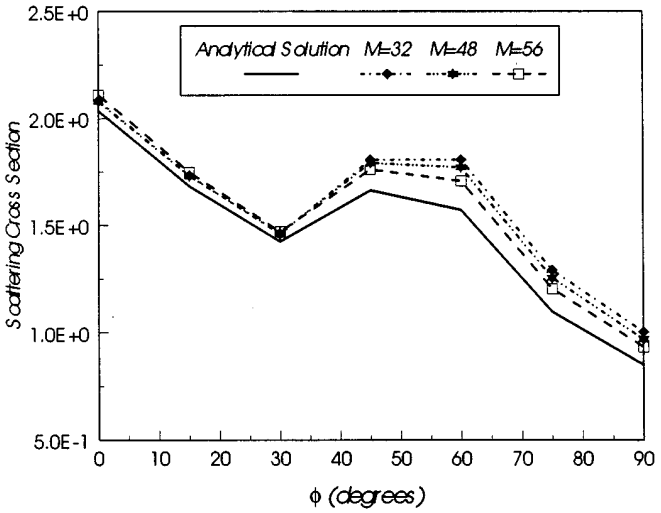


FIG. 9. Convergence of BIE solutions for $h=0.000001a$, shear wave $(K_T a=4)$.

numerical burden is associated with the nearly singular integrals in this approach. This composite BIE is demonstrated to be very effective in the study of scattering from open cracks. Preliminary numerical results show that scattered fields for an open crack with an opening of $2h$ will depart significantly from those fields for a tight crack when h is larger than 5% of the radius of the crack, especially for plane shear waves. All these suggest that the composite BIE formulation is not only robust but also useful in providing information about the physics of thin shape problems. It can fill the gap between the current available one surface models for thin shapes and the real situations or experimental calibrations dealing with open cracks, notches, and rough cracks. The composite BIE formulation is especially valuable for problems in which both the fictitious eigenfrequency difficulty and the thin body breakdown difficulty have to be dealt with.

The composite BIE formulation developed in this paper can be applied to studies of scattering from thin inclusions in materials, thin layer interfaces or interface open (as well as

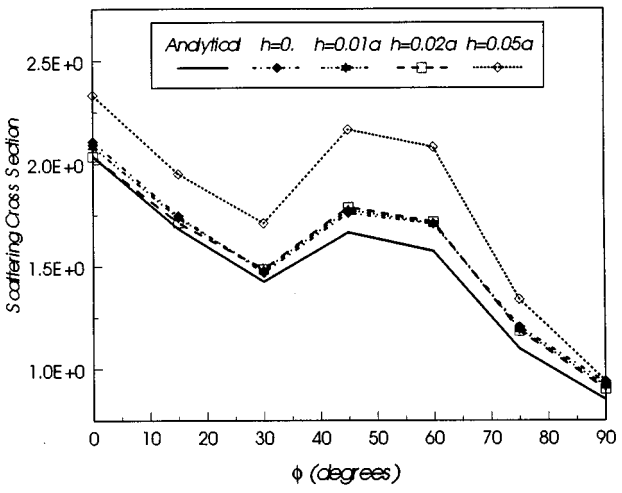


FIG. 10. Solutions using 56 elements at different openings, shear wave $(K_T a=4)$.

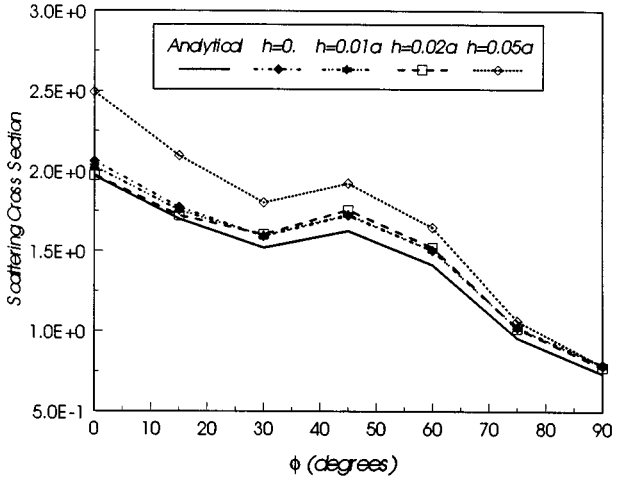


FIG. 11. Solutions using 56 elements at different openings, shear wave $(K_T a=3)$.

closed) cracks, fluid-thin shell like structure interactions, all of which are demanding problems. More complicated nu-merical example problems are being studied along these lines and the results will be reported in future papers.

ACKNOWLEDGMENTS

Part of this work was supported by NIST under cooperative agreement #70NANB9H0916 and was performed at the Center for NDE, Iowa State University. We would like to thank Dr. Ron Roberts at the Center for NDE for helpful discussions. The first author (YJL) would also like to acknowledge the support for his research from the University of Cincinnati.

[1] R. Martinez, ''The thin-shape breakdown (TSB) of the Helmholtz integral equation,'' J. Acoust. Soc. Am. **90**, 2728–2738 (1991).
[2] G. Krishnasamy, F. J. Rizzo, and Y. J. Liu, ''Scattering of acoustic and elastic waves by crack-like objects: the role of hypersingular integral equations,'' in *Review of Progress in Quantitative Nondestructive Evaluation*, Brunswick, Maine (Plenum, New York, 1991), Vol. 11-A.
[3] G. Krishnasamy, F. J. Rizzo, and Y. J. Liu, ''Boundary integral equations for thin bodies,'' Int. J. Numer. Methods Eng. **37**, 107–121 (1994).
[4] C. Y. R. Cheng, A. F. Seybert, and T. W. Wu, ''A multidomain boundary element solution for silencer and muffler performance prediction,'' J. Sound Vib. **151**, 119–129 (1991).
[5] A. F. Seybert, Z. H. Jia, and T. W. Wu, ''Solving knife-edge scattering problems using singular boundary elements,'' J. Acoust. Soc. Am. **91**, 1278–1283 (1992).
[6] T. Terai, ''On calculation of sound fields around three dimensional objects by integral equation methods,'' J. Sound Vib. **69**, 71–100 (1980).
[7] G. Krishnasamy, T. J. Rudolphi, L. W. Schmerr, and F. J. Rizzo, ''Hypersingular boundary integral equations: some applications in acoustic and elastic wave scattering,'' J. Appl. Mech. **57**, 404–414 (1990).
[8] R. Martinez, ''A boundary integral formulation for thin-walled shapes of revolution,'' J. Acoust. Soc. Am. **87**, 523–531 (1990).
[9] T. W. Wu and C. G. 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).
[10] T. W. Wu, ''A direct boundary element method for acoustic radiation and scattering from mixed regular and thin bodies,'' J. Acoust. Soc. Am. **97**, 84–91 (1995).
[11] T. W. Wu and C. G. Wan, ''Muffler performance studies using a direct mixed-body boundary element method and a three-point method for evaluating transmission loss,'' J. Vib. Acoust. **118**, 479–484 (1996).
[12] W. Lin and L. M. Keer, ''Scattering by a planar three-dimensional crack,'' J. Acoust. Soc. Am. **82**, 1442–1448 (1987).

[13] D. E. Budreck and J. D. Achenbach, ''Scattering from three-dimensional planar cracks by the boundary integral equation method,'' J. Appl. Mech. **55**, 405–412 (1988).
[14] G. Krishnasamy, F. J. Rizzo, and Y. J. Liu, ''Some advances in boundary integral methods for wave-scattering from cracks,'' Acta Mech. (Suppl.) **3**, 55–65 (1992).
[15] A. J. Burton and G. F. Miller, ''The application of integral equation methods to the numerical solution of some exterior boundary-value problems,'' Proc. R. Soc. London, Ser. A **323**, 201–210 (1971).
[16] D. S. Jones, ''An exterior problem in elastodynamics,'' Math. Proc. Camb. Philos. Soc. **96**, 173–182 (1984).
[17] D. S. Jones, ''Boundary integrals in elastodynamics,'' IMA J. Appl. Math. **34**, 83–97 (1985).
[18] F. J. Rizzo, D. J. Shippy, and M. Rezayat, ''A boundary integral equation method for radiation and scattering of elastic waves in three dimensions,'' Int. J. Numer. Methods Eng. **21**, 115–129 (1985).
[19] P. J. Schafbuch, F. J. Rizzo, and R. B. Thompson, ''Boundary element method solutions for elastic wave scattering in 3D,'' Int. J. Numer. Methods Eng. **36**, 437–455 (1993).
[20] Y. J. Liu and F. J. Rizzo, ''Hypersingular boundary integral equations for radiation and scattering of elastic waves in three dimensions,'' Comput. Methods Appl. Mech. Eng. **107**, 131–144 (1993).
[21] Y. J. Liu and F. J. Rizzo, ''A weakly-singular form of the hypersingular boundary integral equation applied to 3-D acoustic wave problems,'' Comput. Methods Appl. Mech. Eng. **96**, 271–287 (1992).
[22] Y. J. Liu, D. Zhang, and F. J. Rizzo, ''Nearly singular and hypersingular integrals in the boundary element method,'' in *Boundary Elements XV* (Computational Mechanics, Worcester, MA, 1993), pp. 453–468.
[23] P. A. Martin and G. R. Wickham, ''Diffraction of elastic waves by a penny-shaped crack: analytical and numerical results,'' Proc. R. Soc. London, Ser. A **390**, 91–129 (1983).

Experimental study of sound scattering by elastic spheroids

X. L. Bao
Department of Physics, Catholic University of America, Washington, DC 20064 and Department of Mechanical Engineering, Auburn University, Auburn, Alabama 36849-5341

Herbert Überall
Department of Physics, Catholic University of America, Washington, DC 20064

J. Niemiec
Naval Surface Warfare Center, Carderock Division, West Bethesda, Maryland 20817-5700

(Received 17 May 1993; accepted for publication 14 April 1997)

Experimental results are presented for sound scattering from elastic prolate and oblate spheroidal scattering objects. The backscattered waveform, the target strength versus incident angle, and the frequency response of the echoes are obtained directly from experimental data or via numerical processing of these data. Analytical models and computer codes are developed for estimating the arrival times of the elastic echoes and the resonance modes. © *1997 Acoustical Society of America.* [S0001-4966(97)07007-0]

PACS numbers: 43.40.At, 43.20.Fn [PJR]

INTRODUCTION

The scattering of sound by submerged objects is an important approach towards the exploration of the nature of the object. As is well known, the resonant modes of elastic objects, such as spheres and circular cylinders, can be successfully extracted from the scattered signals. An example is provided in the paper by Flax *et al.*[1] where relations between sound scattering and the resonances of objects are established in terms of the ''Resonance Scattering Theory'' (RST).

For objects of finite size and general shape, numerical methods, especially the *T*-matrix method, are often used to calculate the ''form function'' (i.e., far-field backscattering amplitude) and correspondingly the resonance response. A number of results for prolate spheroids with various aspect ratios were obtained in this way by Werby and Tango[2] and by Hackman *et al.*[3,4]

In papers by Überall *et al.*,[5,6] the phase matching principle for surface waves along closed surface paths was used to estimate the resonance frequencies of finite cylinders or spheroids for the case of axial incidence. For spheroids, the phase velocity of surface waves at any point of the closed path is taken to equal that on a sphere with the same radius of curvature along the direction of propagation. This is referred to as the ''tangent sphere model.'' The predictions of resonance frequencies by this model are in good agreement with the results calculated by *T*-matrix methods, especially at higher values of ka. At lower frequencies, namely where the products kR_1 or kR_2 are small, R_1 and R_2 being longitudinal and transverse radii of curvature along the surface paths, respectively, the estimation errors increase because the surface wave cannot be formed perfectly at low frequency, and the effects of transverse curvature on the phase velocities increase when kR_2 is small. This problem was presented in papers by Hackman *et al.*,[3,4] and a ''bar wave'' approach was used for cylinders with high aspect ratio. In our earlier paper,[7] results of a modified phase matching method, using an ''infinite cylinder model'' on the cylindrical portion of hemispherically end-capped cylinders, are compared with those of the tangent sphere model.

Experimental studies have also been concerned with sound scattering and resonance predictions for finite elastic cylinders. Experimental methods were developed in order to determine the character of the scattering objects, where the incident signals used can be cw (long pulses) as in the method of Maze and Ripoche,[8] or short pulses as in the method of Numrich *et al.*[9] and de Billy.[10]

In the present paper, experimental results of sound scattering are presented for prolate and oblate elastic spheroids. The backscattered waveform, the target strength versus incident angles, and the frequency response of the echoes are obtained directly from experimental data or via numerical processing of these data. In our experiment, short pulses are incident in the axial direction and also in discrete oblique directions every 5° from the spheroid axis.

Corresponding analytical methods, based on the picture of circumferential surface waves, are presented also. The special surface paths along which the sound waves can be returned are searched for and the time delay of the first elastic echo, related to specular reflection, is estimated as a function of incident angle. For estimating the resonance modes, the phase-matching principle of surface waves is extended to the case of oblique incidence. The tangent sphere model must basically be applicable for estimating the phase velocities in our experiment, because the values of kR_1 or kR_2 are not too small here. Although the possible resonance modes can be approximately estimated by the phase-matching method, it is difficult to precisely analyze the actual resonance responses obtained in our experiment for relatively higher values of ka. Further improvements are needed in our approach to be able to deal with these aspects of the problem.

TABLE I. Type and size of steel spheroids and range of frequency spectrum, expressed by ka and kb.

Type	a/b	a (mm)	b (mm)	ka	kb
Prolate	1/5	7.0	35.0	13–22	66–110
Prolate	1/2	9.5	19.0	18–30	36–60
Oblate	2/1	15.2	7.6	29–48	14–24
Oblate	5/1	20.5	4.1	39–64	7.7–13

I. EXPERIMENT AND DATA PROCESSING

A. Experiment

The experiments on underwater sound scattering were carried out in a small tank of dimensions 1.2 by 0.45 by 0.5 m. The type and size of the steel spheroids are listed in Table I.

The electric signal formed by the generator is a short sinusoidal 3-cycle pulse of 600 kHz, but the waveform of the direct sound signal from projector to receiver is slightly distorted by the transducers. The range of the frequency spectrum of the direct signal, expressed by ka and kb, is also listed in Table I for the objects with various aspect ratios a/b.

Two transducers of 3-cm diameter are placed in the water at the same depth and close to each other (at a bistatic angle of less than 3°), being used as projector and receiver, respectively. The short pulses formed by the waveform generator are amplified by a power amplifier, then are sent to a projector. The sound pulses emitted by the projector are scattered by the object submersed in water. The backscattered waves are picked up by the hydrophone, and are amplified and digitized by a digitizer. Finally, data are saved on disks for further processing.

B. Data processing

Some of our experimental results of backscattered signals for spheroids are shown in Fig. 1. The data processing is done on a VAX computer, and some special codes were developed for it.

1. Spectrum analysis

(a) The amplitude spectrum of the backscattered waves can be obtained by using fast Fourier transform (FFT) technology directly. Suppose $e(t)$ and $E(ka)$ are the echo waveform and its Fourier transform, respectively. Let

$$E(ka)=E_1(ka)+iE_2(ka), \tag{1}$$

then the amplitude spectrum can be obtained by

$$E_A(ka)=[E_1^2(ka)+E_2^2(ka)]^{1/2}, \tag{2}$$

and the normalized amplitude spectrum (''form function'') can be expressed as

$$A(ka)=E_A(ka)/D_A(ka), \tag{3}$$

where $D_A(ka)$ is the amplitude spectrum of the direct signal.

(b) The resonance spectrum can be extracted from the backscattered waves by using the method shown in Refs. 9, 10. Here, the FFT is used on the waveform from which the specularly reflected wave has been removed. This is different from the extraction of the amplitude spectrum (the form function), obtained by FFT without removing the specular wave.

2. Target strength (TS)

The concept of TS is used in acoustics to obtain the relative echo intensity of submerged objects. In the present paper, two distinct data processing methods are employed for obtaining the target strength.

(a) In the first method, corresponding to the geometrical-acoustics approximation, only the amplitude of the specular reflection pulse is considered. For steel objects at higher frequencies, the amplitude of the specular reflection wave is nearly the same as the steady-state amplitude for rigid objects or for elastic objects which are not in a state of resonance. Let $p(\beta)$ be the amplitude of the specular reflection pulses at the incident angle β; then the target strength (TS) is obtained from[11]

$$\mathrm{TS}_p(\beta)=20\log[p(\beta)], \tag{4}$$

where $\mathrm{TS}_p(\beta)$ is a relative value that differs from the real target strength by a constant factor.

(b) In the second method used to estimate TS (β), the total energy of the (entire) echo signal, and also the (partial) energy of the elastic echo signal at the incident angle β is considered, respectively. If $p(t,\beta)$ is selected corresponding to such an echo signal, then the relative strength $\mathrm{TS}_E(\beta)$ can be obtained by

$$E(\beta)=\int p^2(t,\beta)dt, \tag{5}$$

$$\mathrm{TS}_E(\beta)=10\log[E(\beta)]. \tag{6}$$

Because the total energy including the elastic echoes is being considered here, it is expected that $\mathrm{TS}_E(\beta)$ will be affected by the elasticity of the objects.

The result of this TS versus incidence angle are shown in Fig. 2. As the pulse signal used has a center frequency (600 kHz) and is spread over a certain frequency band, the TS represents an average around the center frequency. For a TS referring to a single frequency, it is better to use the form function or the resonance spectrum in place of $p(\beta)$ or $E(\beta)$.

II. METHOD OF NUMERICAL ESTIMATION

A. Time sequences of backscattered pulses and surface wave paths

The sound paths of echoes generally include geometrical reflection and surface circumferential paths. In the following, the surface paths on spheroids are discussed.

1. Axial incidence

This is the simplest case and has been discussed in previous work. Only the surface paths along the meridians exist here, and the surface waves of a given type propagating along all of these paths add up in phase at receiver, because all of the meridian paths have the same shape and path length. For estimating the time sequences, it is necessary to know the phase velocities which are varying along the sur-

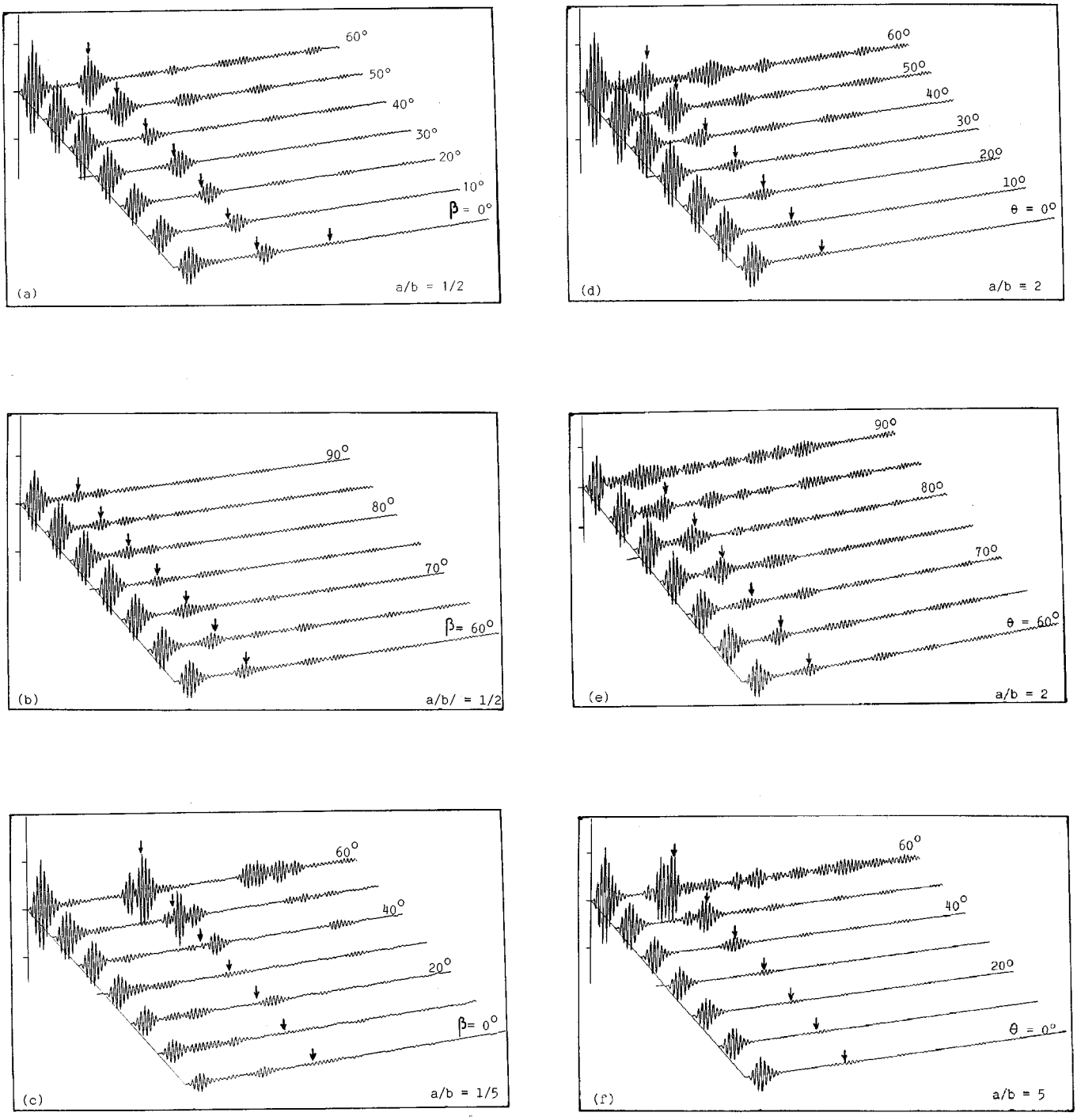


FIG. 1. Echo waveform versus incident angles (3-D). (a) Prolate spheroid, $a/b=1/2$, $\beta=0°-60°$; (b) Same as (a), $\beta=60°-90°$; (c) Prolate spheroid, $a/b=1/5$, $\beta=0°-60°$; (d) Oblate spheroid, $a/b=2$, $\theta=0°-60°$; (e) Same as (d), $\theta=60°-90°$; (f) Oblate spheroid, $a/b=5$, $\theta=0°-60°$.

face path. But for objects such as spheroids, it is difficult to solve for these analytically. Some approximate methods are presented for estimating the phase velocity. In this paper, the ''tangent sphere model'' is employed.

2. Broadside incidence

For an infinite circular cylinder, the surface paths are a set of closed circles and are independent of the incident position along the cylinder in the case of broadside incidence. But for a finite rotationally symmetric object such as the spheroid, the surface paths are no longer closed, and are dependent on the position where the sound waves are incident. A circular path is formed only at the center position of the spheroid. This path is also a principal path along which the energy of sound waves returns to the receiver. The paths become helicoidal when the points of incidence are off center. (A virtual source of scattered rays could possibly be found by considering the paths of surface waves in the vicinity of the principal path. The vertically scattered waves are reradiated as if coming from this virtual source.) The amplitude of the echoes will be affected by this. It is expected that the amplitude of echoes increases when the

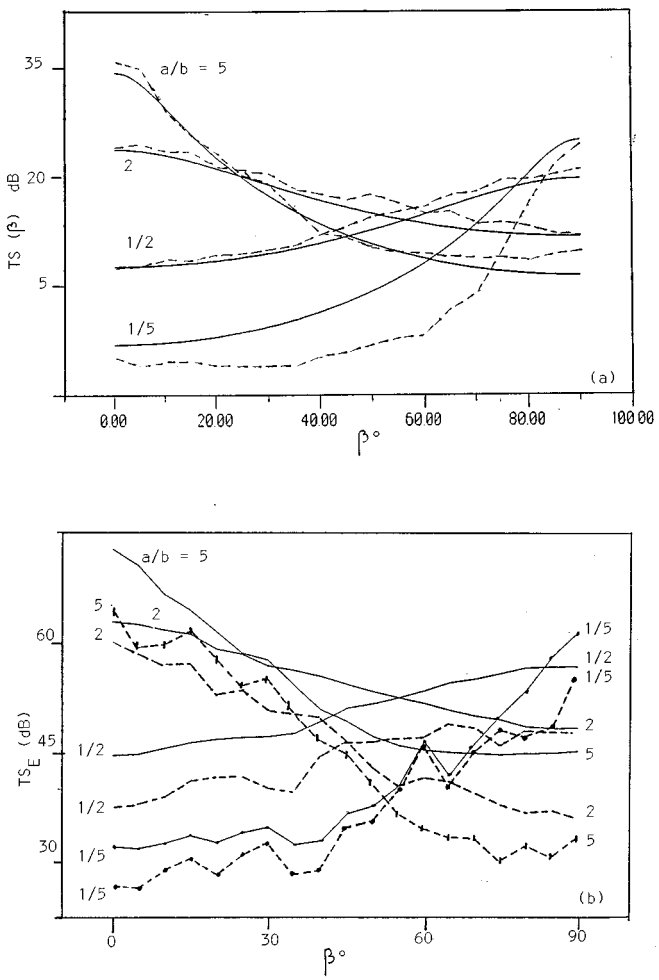


FIG. 2. Target strength according to theoretical formula and experimental data processing. (a) Solid lines: formula of geometrical acoustics; dashed lines: in terms of specular reflection pulses. (b) Solid lines: in terms of total energy of echoes; dashed lines: in terms of energy of elastic waves.

length of the prolate spheroids increases in the axial direction. In the case of broadside incidence, a meridian path can also be formed when the points of incidence are in the axial plane; it changes when the incident points are away from the axial plane. The path changes with changes of incident position are more sensitive for prolate spheroids of larger length, and are stable for oblate spheroids. Therefore, it can be expected that the echoes which are related to meridian paths will be stronger for oblate spheroids than for prolate spheroids.

3. Oblique incidence

In the case of oblique incidence, the paths of surface waves are helicoidal lines with distinct start angles, and the sound scattering fields are dependent on the paths. Only around a few special paths can the sound waves return to the direction of the receiver. The meridian paths always exist for oblique incidence but they play an important role in the echoes only when the incident direction is near-axial. For certain spheroids and certain incident angles, the principal surface paths can be found. For example, the paths for prolate spheroids as shown in Fig. 3, lead to a sound-wave return while passing through the backside once. For a certain incident

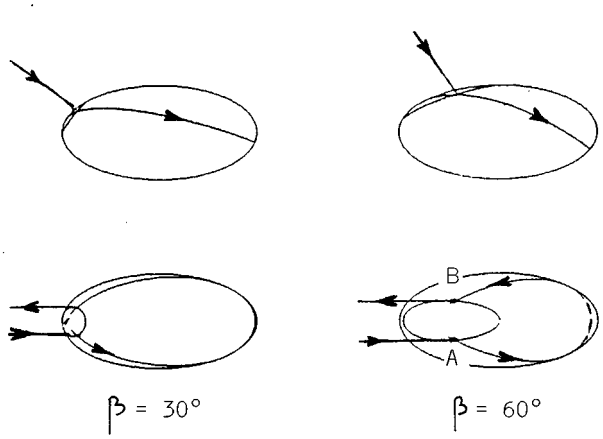


FIG. 3. Example of special surface paths.

angle, it can be found that a point A exists on the ''line of entry'' (indicated in Fig. 3) where the surface path starting from A leads back to a special point B which is symmetric to A on the line of entry with respect to the incident plane through the semimajor axis, and the surface path on both sides of the incident plane also shows symmetry; therefore the return conditions are satisfied.

A computer program was developed to find these special surface paths. The estimated propagation time, in terms of special surface paths, versus incident angles are listed in Table II.

For estimating the time sequences of the elastic echoes in terms of the surface paths, the group velocity along a given surface path must be estimated beforehand. In Fig. 4, the phase velocity (c_p) and group velocity (c_g) versus normalized frequency ka are plotted for a solid steel sphere. These quantities c_p and c_g vs ka were calculated by first solving the characteristic equation for the resonance values $(ka)^*_{nl}$ for the various waves labeled by integers l, then finding c_p from $c_p/c_0=(ka)^*/(n+1/2)$, see Ref. 12. The group velocities are relatively stable compared with the phase velocities for the surface wave mode types $l=1,2,3$. In this paper, for simplicity, constant group velocity values are used for estimation purposes. In Fig. 5, the estimated arrival times versus incident angles obtained in this way for type $l=1$ surface waves are compared with experimental results.

TABLE II. Surface-wave time delays relative to the specular-reflection echo as estimated for the first elastic echo pulses.

Incident angles β (°)	$a/b=1/5$ t (μs)	$a/b=1/2$ t (μs)	$a/b=2/1$ t (μs)	$a/b=5/1$ t (μs)
5	49.4	30.2	20.5	20.0
15	49.3	30.0	21.6	24.1
30	48.9	29.2	24.2	31.6
45	47.9	27.7	26.7	35.9
60	44.7	25.1	28.3	37.4
75	34.9	21.4	29.4	37.6
85	22.0	18.9	29.9	37.6
89	16.4	18.4	-	-

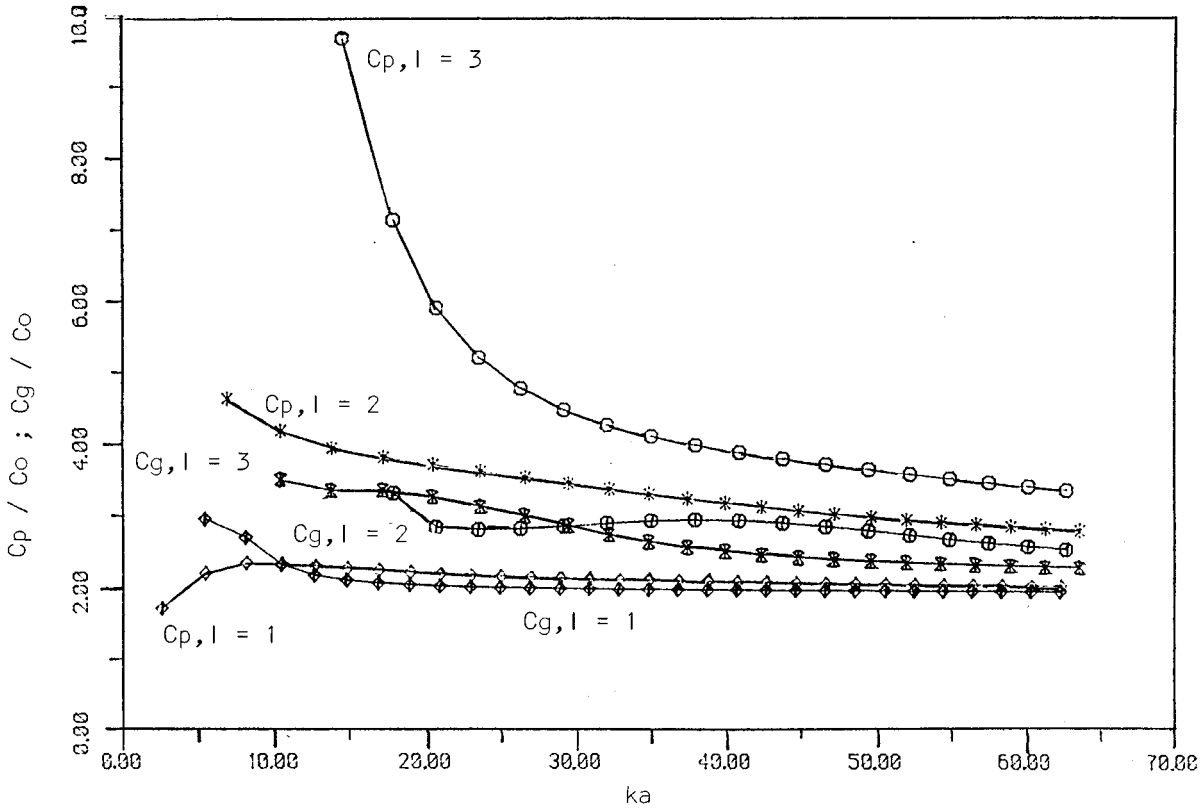


FIG. 4. Phase velocity and group velocity versus ka for steel sphere ($c_l = 5790$ m/s, $c_t = 3100$ m/s, $c_0 = 1480$ m/s).

B. Resonances estimated by the phase-matching method

For the case of axial incidence, a resonance condition based on the phase-matching principle has been formulated by Überall *et al.*,[5] and was used to estimate the resonances of elastic prolate spheroids and cylinders with hemispherical end caps. It reads

$$\oint k_1 \, ds = 2(n+1/2)\pi, \tag{7}$$

where n is an integer, and the term 1/2 arises from phase jumps when the surface waves pass through focal points. The quantity k_1 is the propagation constant for the lth mode of surface waves and it varies along the path as the surface curvature varies.

In a paper by Merchant *et al.*,[6] the phase-matching principle is used to estimate the eigenfrequencies of conducting spheroids for the case of oblique incidence. The surface path follows a helicoidal geodesic (shortest surface distance between two points), and closes into itself. Following Ref. 6, the resonance modes of steel spheroids are estimated in the present study. It was found that sets of redundant modes are generated by this method and it is not easy to eliminate these. In the following, another method, also based on the phase matching principle, but not necessarily on closed paths, is presented and compared with the former method.

1. Phase matching of a surface wave along closed surface paths

In the following, the main relations are listed according to Ref. 6. The propagation of surface waves along geodesics on prolate spheroids has previously been discussed by Burnside *et al.*[13] Geodesics are defined as curves of zero geodesic curvature. The general geodesic equation[6,14]

$$\phi(\theta_f) - \phi(\theta_0) = C \int_{\theta_0}^{\theta_f} \frac{[\rho^2(\theta) + (d\rho/d\theta)^2]^{1/2}}{\rho(\theta)\sin\,\theta[\rho^2(\theta)\sin^2\,\theta - C^2]^{1/2}}\, d\theta, \tag{8}$$

is valid for any surface of revolution. Here, spherical coordinates ρ, θ, ϕ are used. Equation (8) expresses the difference in azimuth between two points 0, f on a geodesic in terms of an integral between the two associated polar angles θ_0 and θ_f, with C given by

$$C = p(\theta_0)\sin\,\theta_0\,\sin\,\alpha_0, \tag{9}$$

α_0 being the starting angle of the geodesic at point 0.

For a spheroid with equatorial radius a and polar half-axis b, the starting point is placed at the equator [$\theta_0 = \pi/2$, $\phi(\theta_0) = \phi_0 = 0$]. The geodesic path on a spheroid is then found to be given by

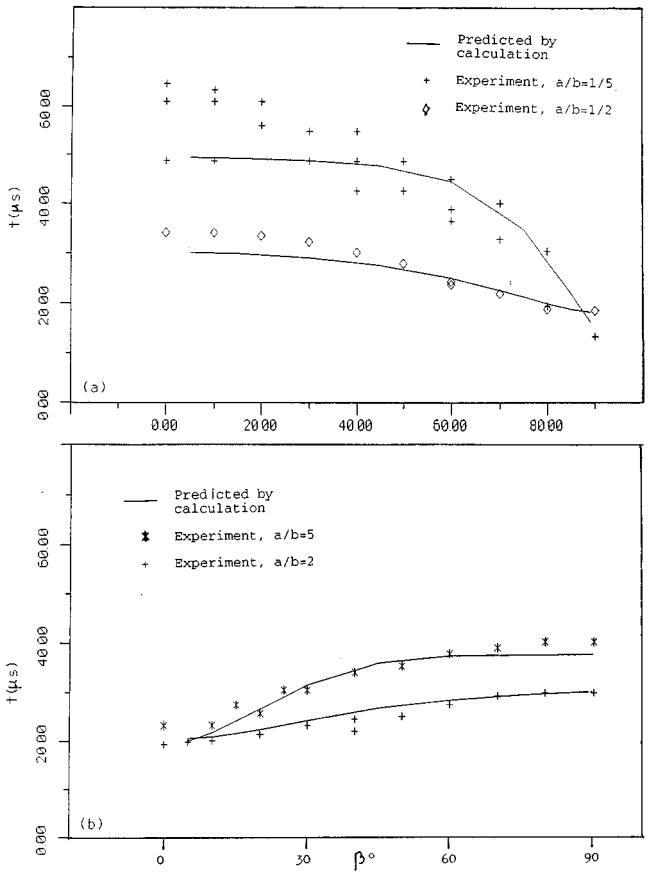


FIG. 5. Arrival time of the first elastic echo versus incident angles (a) for prolate spheroids; (b) for oblate spheroids.

$$\phi(\theta_f)=\int_{\pi/2}^{\theta_f}\frac{d\theta}{\sin^2\theta}\times\left[\frac{1+\eta^{-4}\cot^2\theta}{(\cot^2\alpha_0-\eta^{-2}\cot^2\theta)(1+\eta^{-2}\cot^2\theta)}\right]^{1/2}, \tag{10}$$

where $\eta=b/a$. The turning points (points of minimum polar angle) are in this case determined by

$$\cot\theta_m=\pm\eta\cot\alpha_0. \tag{11}$$

To close a geodesic, one must require that

$$4\phi_m N=2\pi K, \tag{12}$$

where N and K are integers. Closure thus take place after N cycles of the geodesic, which are fitted into K azimuthal circuits around the spheroid circumference.

For the propagation along a closed helicoidal geodesic, with a starting angle α_0 at the equator, the expressions according to phase matching are, with $\nu_1\equiv k_1(\theta)R(\theta)$, k_1 being the local propagation constant:

$$J\equiv\frac{2N}{\pi}\int_{\pi/2}^{\theta_m}\frac{\nu_1[k^*R(\theta)]}{R(\theta)}\frac{dS(\theta)}{d\theta}\,d\theta=M+\frac{N}{2}, \tag{13}$$

$$R(\theta)=\frac{a}{b(a^2\cos^2\alpha_0+b^2\sin^2\alpha_0)}\times\frac{(a^4\cos^2\theta+b^4\sin^2\theta)^{3/2}}{(a^2\cos^2\theta+b^2\sin^2\theta)^{3/2}}, \tag{14}$$

$$\frac{dS(\theta)}{R(\theta)}=\frac{b^3\sin\theta d\theta}{a^4\cos^2\theta+b^4\sin^2\theta}\times\frac{a^2\cos^2\alpha_0+b^2\sin^2\alpha_0}{(b^2\sin^2\theta\cos^2\alpha_0-a^2\cos^2\theta\sin^2\alpha_0)^{1/2}}. \tag{15}$$

2. Phase-matching conditions satisfied along longitudinal and latitudinal paths

The wave-number vectors k_s along the geodesic can be decomposed into two components, one along a longitudinal line (meridian) and the other one along a latitudinal line (circumferential line). To obtain phase matching, the necessary conditions can be written as

$$\int_{\varphi=0}^{2\pi}k_\varphi\,ds_\varphi=2\pi m, \tag{16}$$

$$2\int_{\theta=0}^{\pi}k_\theta\,dS_\theta=2\pi(n+1/2), \tag{17}$$

or they can be simplified if the wave number is not dependent on surface curvature:

$$\int_0^{2\pi}k_s(\theta)\sin\alpha(\theta)r(\theta)d\varphi=2\pi ak_s(\theta_0)\sin\alpha(\theta_0)=2\pi m, \tag{18}$$

$$4\int_0^{\pi/2}k_s(\theta)\cos^2\alpha(\theta)\frac{dS_\theta}{d\theta}\,d\theta=2\pi(n+1/2). \tag{19}$$

3. Comparison of the aforementioned methods

The closure of paths is required by the first method, which seems intuitively correct because it is an immediate extension from the sphere, or from an infinite cylinder in the case of broadside incidence. But the closure condition is not actually a basic condition for resonances of general objects.

For example, consider an infinite circular cylinder insonified by an obliquely incident sound wave. In that case the surface paths are sets of helical lines and the closure condition is not satisfied. But the phase-matching condition can be satisfied circumferentially. Although the sound paths cannot be closed, the wave fronts overlap in phase with each other while the resonance condition is satisfied.

According to the second method the phase-matching conditions are satisfied in two directions simultaneously. This ensures that wavefronts can overlap in phase with each other while the surface waves propagate helicoidally around the objects. Returning to the first method, although the phase-matching condition is satisfied along the closed paths, this cannot ensure that the phase-matching conditions are satisfied simultaneously along the circle and the meridian. The fact is that the latter is usually not satisfied, and this is also the reason that redundant modes are generated.

C. Estimation of the target strength

The general formula for estimating the target strength of convex finite objects with a smooth surface is presented by Urick.[11] It is based on the geometrical-acoustic approximation at higher frequencies. The contributions due to the elastic nature of the objects are ignored. In the formula, only the principal radii of curvature at the points of specular reflection enter:

$$\mathrm{TS}(\beta) = 10 \log(R_1 R_2/4); \tag{20}$$

here R_1 and R_2 are the principal radii of curvature. For spheroids, the corresponding expressions are

$$R_1(\beta) = a^2 R_2^3/b^4, \tag{21a}$$

$$R_2(\beta) = (a^4 x_1^2/b^4 + z_1^2)^{1/2}, \tag{21b}$$

$$x_1 = b\ \cos\ \beta/b(\sin^2\ \beta/b^2 + \cos^2\ \beta/a^2)^{1/2}, \tag{21c}$$

$$z_1 = a(1 - x_1^2/b^2)^{1/2}, \tag{21d}$$

where x_1, z_1 are the coordinates of the points of specular reflection, and a and b are the semiminor axis and semimajor axis, respectively.

III. RESULTS AND DISCUSSION

As shown above, a set of results is obtained in our experiments for spheroids. In the following, the results are discussed, and are compared with the results of the numerical estimation.

A. Target strength (TS) and its dependence on incident angles

(i) The dynamic range of target strength variation with incident angles is dependent on the ratio a/b. As shown in Fig. 2(a) and (b), for oblate spheroids, TS reaches its maximum value for axial incidence and decreases while the incidence angles increase. The dynamic range of TS increases with increasing ratio a/b. Conversely, for prolate spheroids, TS reaches its maximum for broadside incidence and the dynamic range increases when a/b decreases.

(ii) The results for TS versus incidence angles, obtained from the specular reflection of waves, are in general similar to the results predicted by the surface curvatures at the specular reflection points [Fig. 2(a)]. The results are as expected for the specular reflected waves. But for the TS obtained from the total echo energy [Fig. 2(b)], because it also includes elastic echoes, the effects of resonances must be considered. The experimental results show that the effects of elastic echoes can be found at several incident directions, but generally, the results for TS do not differ greatly from the specularly reflected wave results.

(iii) The results obtained for TS that are based on the energy of elastic echoes alone are obviously different from the results obtained from the total energy of echoes [Fig. 2(b)]. In this case, TS fluctuates when the incident angles are changed. The fluctuation can be understood because the possible resonance modes and their amplitude of excitation by sound waves are dependent on the incident angles. The bottom line is that the total echo always contains the specular, and thus maintains some semblance of a TS floor.

TABLE III. The average group velocities c_g of surface waves estimated by the arrival time t_1 of the first elastic echo (d_s, d_w are distance along surface paths and in water, respectively; c_w is sound speed in water. The spread in the c_g is due to the experimental results of t_1 being dispersed).

a/b	t_1 (μs)	d_s (mm)	d_w (mm)	c_g (m/s)	c_g/c_w
0.2	48.8–61.0	145.5	0.46	2397–3000	2.03–1.62
0.5	34.2	86.5	1.49	2610	1.76
2.0	19.3	49.8	5.25	3168	2.14
5.0	23.2	46.7	5.52	2404	1.62

B. Discussion concerning the time sequences of echo pulses

For the case of axial incidence, the average group velocities of surface waves, estimated according to the arrival time of the first strong elastic echo and specification of the surface paths, are listed in Table III. The average velocities are in a range from 1.6 to 2.0, which is slightly less than the expected value from using the dispersion relations of elastic spheres.

As shown in the figures of the waveform (Fig. 1), the time sequences of the echo pulses vary with incidence angles. In the following, the arrival times of the first strong elastic echoes are obtained from experimental records of the waveform, and are compared with numerical estimations obtained in terms of special surface paths.

As shown in Figs. 5 and 1, the arrival times of the first strong elastic echo pulses vary with incident angles and can be basically explained by the surface waves traveling along the special surface paths.

For prolate spheroids with higher aspect ratio ($b/a = 5$), the difference between the results of experiment and calculation increases when the incident angles decrease to near axial incidence. The first strong elastic echoes seem to correspond to an average velocity lower than the expected value for the elastic surface waves, but this cannot be explained either by inside-reflection paths or by creeping waves. We could, however, attribute this to the fact that the experimental results for the 5:1 spheroids are not as good as for the 2:1 spheroids; the elastic echoes there are quite weak for axial incidence.

C. Frequency responses of backscattered waves

As mentioned above, the frequency responses, including the amplitude spectrum and resonance spectrum, for backscattered signals are obtained from experiment by data processing and the possible resonance modes are estimated according to the phase matching principle.

(i) As shown in the figures of the amplitude spectrum (resonance spectrum), Fig. 6, a set of resonance peaks can be discerned. As expected, the situation of the resonance spectrum in relation to frequencies and to incidence angles is rather complicated because the sound paths are complex for solid elastic spheroidal objects.

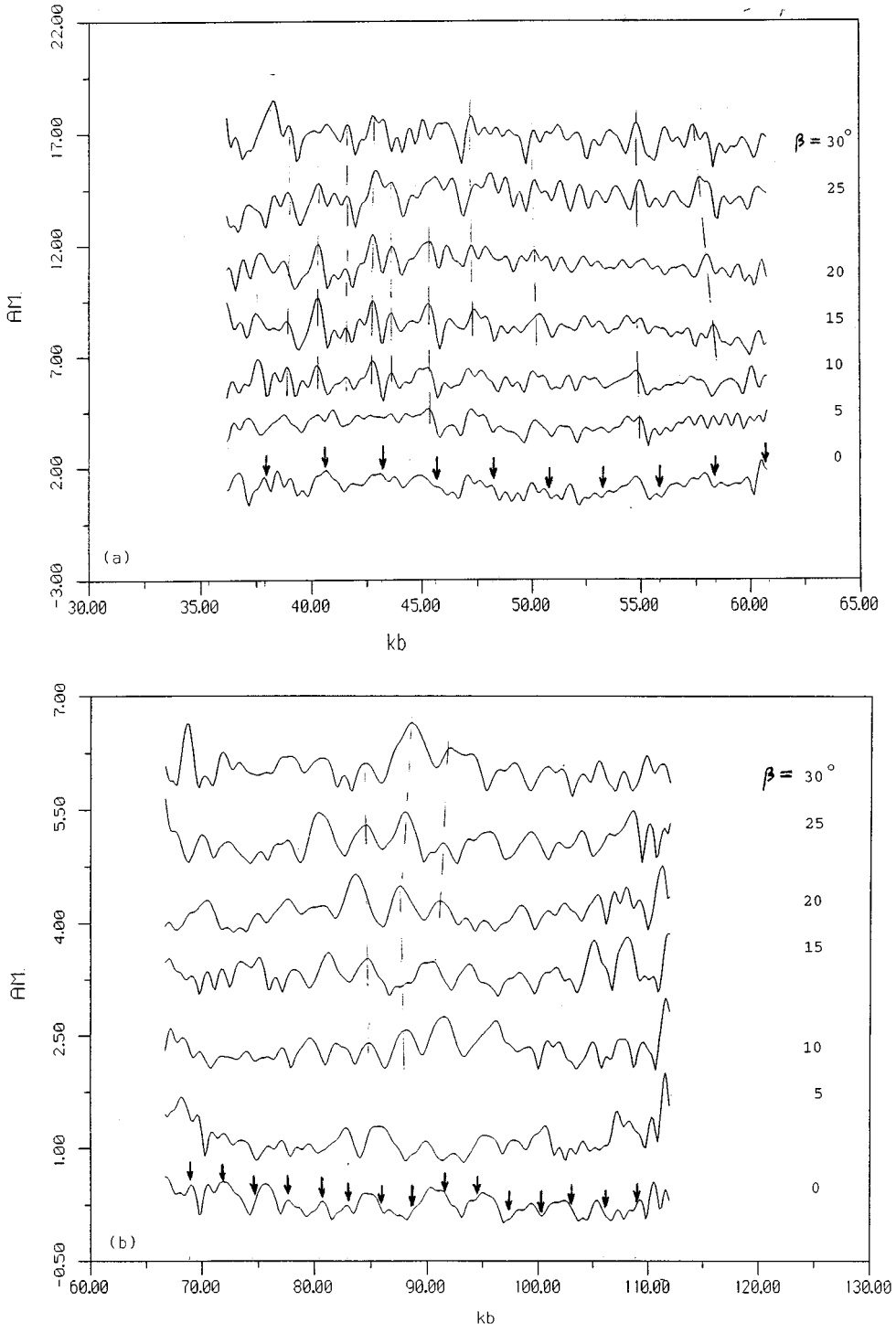


FIG. 6. Resonance responses by experimental data processing, for prolate spheroid (a) $a/b=1/2$; (b) $a/b=1/5$.

The distribution of peaks is similar for prolate spheroids while the incidence angles are small ($\beta=0°-30°$) and is similar for oblate spheroids near broadside incidence ($\beta=60°-90°$). This can be explained in terms of corresponding time waveforms. As shown in Fig. 1, the time sequences are stable at small incidence angles for prolate spheroids, and are stable for incidence angles near broadside incidence for oblate spheroids.

(ii) The possible resonance modes according to the numerical prediction are calculated for $l=1,2,3$ type surface waves. An example is shown in Fig. 7 for a prolate spheroid ($a/b=1/2$, $l=1$). The average phase velocity, in terms of average curvature along the path and dispersion of a steel sphere, is used in the phase-matching formula. Here the values (m,n) express the mode number along meridian and circumferential directions, respectively. For a given mode (m,n) the starting angle of the helical surface path designated by $\alpha_{0m,n}$ is determined. But the possible starting angles α_0 are also related to the incidence angles β. In Fig. 8, the maximum values $\alpha_{0\,\max}$ are plotted versus incidence

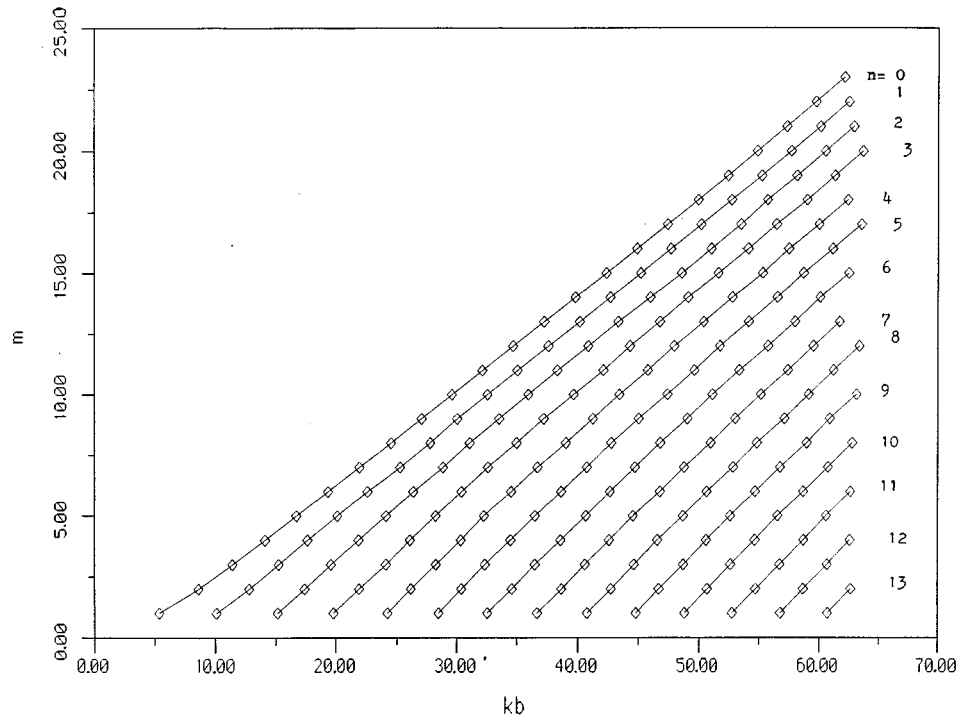


FIG. 7. Possible resonance modes estimated by phase-matching method for $a/b=1/2$ prolate spheroid and for $l=1$ type surface waves.

angles β for different spheroids. In the case of axial incidence only the $(m,0)$ modes can be excited by the incident wave. Further possible modes $(n>0)$, satisfying $\alpha_{0m,n} \leqslant \alpha_{0\max}$, can be excited while the incidence angle β increases.

(iii) For comparison, in Fig. 6, the predicted resonance modes $(m,0)$ for the case of axial incidence are marked by a down arrow. As can be seen from these figures, the predicted modes do not always coincide with the experimental resonance peaks. Because the experimental results of the resonance spectrum are complex as shown, it appears that one cannot arrive at a universal recognition of resonances by using the present simplified theoretical model. For recognition, it is also important to estimate the relative resonance strength

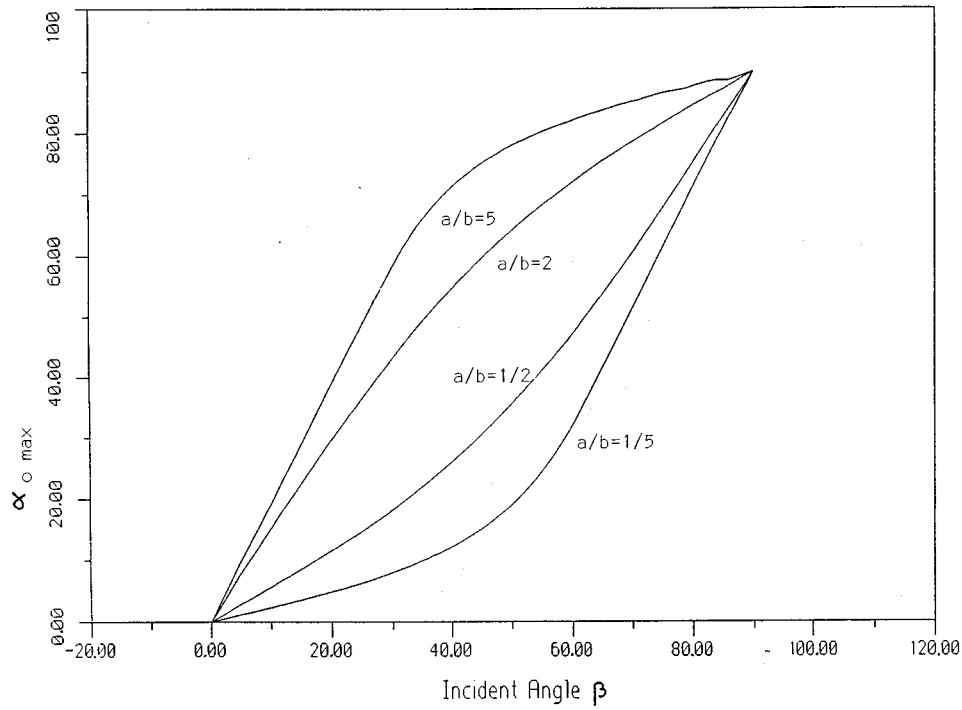


FIG. 8. Maximum starting helical angles (at the equator) versus incidence angles β.

for the possible modes, so that the dominant modes can be found. There are several problems here that need to be considered in further studies.

IV. CONCLUSIONS

(a) An experimental study of sound scattering by elastic spheroids has been carried out. In this paper, a set of interesting results obtained by experiments and by data processing is provided. The results include average properties, time responses, and frequency responses of backscattered waves for objects such as oblate spheroids and prolate spheroids, excited by sound waves with distinct incidence angles.

(b) For explaining the results obtained by the experiments, several theoretical models and codes are developed, and a set of results is provided by numerical calculation. These include mainly the models for estimating the target strength, the time sequences in terms of the special surface paths, and the resonance modes according to the phase-matching principle, besides other topics.

(c) The sound scattering is affected by the elasticity of objects; this can not only be observed in an obvious fashion from the echo waveform and from frequency responses, but it can be observed from the echo strength. As mentioned, the target strength obtained from an energy integration fluctuated with the incidence angles, thus it is obviously different from the results obtained by only considering the specular reflections.

(d) As mentioned, the arrival times of the first elastic echoes coincide basically with the theoretical predictions obtained by considering the surface paths. This shows that the surface waves play an important role in the echo responses for spheroidal objects used in our experiments even in the case of oblique incidence. Therefore, methods based on the theory of surface waves should be suitable for studying the problems related to sound scattering by finite objects, such as spheroids.

(e) Resonance scattering is an important topic for target recognition. In this paper, various relevant results both from experiments and from theoretical models are provided and discussed. But because the problems are rather complicated for finite elastic objects of noncanonical shapes, further research should be carried out.

ACKNOWLEDGMENTS

The authors wish to express their appreciation to Dr. J. Clark for providing facilities for research and experiment and for his advice which made this research possible.

[1] L. Flax, L. R. Dragonette, and H. Überall, ''Theory of elastic resonance excitation by sound scattering,'' J. Acoust. Soc. Am. **63**, 723–731 (1978); see also *Acoustic Resonance Scattering*, edited by H. Überall (Gordon and Breach, New York, 1992).

[2] M. F. Werby and G. J. Tango, ''Numerical study of material properties of submerged elastic objects using resonance response,'' J. Acoust. Soc. Am. **79**, 1260–1268 (1986).

[3] R. H. Hackman and D. G. Todoroff, ''An application of the spheroidal-coordinate-based transition matrix: The acoustic scattering from high aspect ratio solids,'' J. Acoust. Soc. Am. **78**, 1059–1071 (1985).

[4] R. H. Hackman, G. S. Sammelmann, K. L. Williams, and D. H. Trivett, ''A reanalysis of the acoustic scattering from elastic spheroids,'' J. Acoust. Soc. Am. **83**, 1255–1266 (1988).

[5] H. Überall, Y. J. Stoyanov, A. Nagl, M. F. Werby, S. H. Brown, J. W. Dickey, S. K. Numrich, and J. M. D'Archangelo, ''Resonance spectra of elongated elastic objects,'' J. Acoust. Soc. Am. **81**, 312–316 (1987).

[6] B. L. Merchant, A. Nagl, and H. Überall, ''A method for calculating eigenfrequencies of arbitrarily shaped convex targets: Eigenfrequencies of conducting spheroids and their relation to helicoidal surface wave paths,'' IEEE Trans. Antennas Propag. **37**, 629–634 (1989).

[7] X. L. Bao, H. Überall, and J. Niemiec, ''Experiments on excitation of resonances of elastic objects by acoustic pulses and theoretical interpretation,'' Appl. Mech. Rev. **44**, S16–S19 (1991).

[8] G. Maze and J. Ripoche, ''Methode d'isolement et d'identification des resonances (M.I.I.R.) de cylindres et de tubes soumis a une onde acoustique plane dans l'eau,'' Rev. Phys. Appl. **18**, 319–326 (1983).

[9] S. K. Numrich, N. Dale, and L. R. Dragonette, ''Generation and exploitation of plate waves in submerged, air-filled shells,'' in *Advances in Fluid-Structure Interactions* (ASME, New York, 1984), PVP Vol. 78/AMD Vol. 64, pp. 59–74.

[10] M. de Billy, ''Determination of the resonance spectrum of elastic bodies via the use of short pulses and Fourier transform theory,'' J. Acoust. Soc. Am. **79**, 219–221 (1986).

[11] R. J. Urick, *Principles of Underwater Sound* (McGraw-Hill, New York, 1983).

[12] H. Überall, L. R. Dragonette, and L. Flax, ''Relation between creeping waves and normal modes of vibration of a curved body,'' J. Acoust. Soc. Am. **61**, 711–715 (1977).

[13] N. Wang, W. D. Burnside, and R. G. Kouyoumjian, ''A uniform GTD solution for the radiation from sources on a convex surface,'' IEEE Trans. Antennas Propag. **AP-29**, 609–622 (1981).

[14] D. J. Struik, *Differential Geometry* (Addison-Wesley, Reading, MA, 1961), 2nd ed.

Energy flow analysis for curved beams

A. Le Bot, M. N. Ichchou, and L. Jezequel
Laboratoire de Tribologie et Dynamique des Systèmes, UMR CNRS 5513, École Centrale de Lyon 36, avenue Guy de Collongues, BP163, 69131 Ecully cedex France

(Received 25 July 1995; revised 4 November 1996; accepted 31 March 1997)

This paper presents an energy model for the medium- and high-frequency analysis of Love–Kirchhoff curved beams. This model introduced by Nefske and Sung [Statistical Energy Analysis NCA **3**, 47–54 (1987)] for straight beams and investigated further by other authors, is developed for curved rods (tangential or longitudinal waves), and then for curved beams (radial or flexural waves). The exact-energy solution for curved rods or beams is shown to consist of a smooth spatial variation, which the energy model represents, and a spatially oscillating solution, which can be represented by an energy envelope. Finally, a complete energy model is proposed for curved components including both longitudinal and flexural waves. Boundary conditions are also given in this paper. It is shown that this method, which is numerically attractive in the mid- and high-frequency range, predicts the arithmetic mean value of the energy variables. © *1997 Acoustical Society of America.* [S0001-4966(97)05807-4]

PACS numbers: 43.40.Cw [CBB]

LIST OF SYMBOLS

ω	circular frequency
E_0	Young's modulus
E	complex modulus
η	damping ratio
s	curvilinear coordinate
R	radius of curvature
S	area of section
I	inertia
ρ	mass density
N	tensile load
T	shear force
M	bending moment
u	tangential displacement
v	radial displacement
θ	rotation of section
W	energy density
P	active energy flow
p_{diss}	power density being dissipated
p_{lf}	exchanged power density from longitudinal to flexural form
k	complex wave number of curved system
k_∞	complex wave number of straight system
c_g	group velocity
a^+, a^-, b^+, b^-	displacement magnitudes
A^+, A^-	energy magnitudes

INTRODUCTION

In designing structures, one of the main questions is how to predict and control noise and vibration? At low frequencies, several tools allow the vibration level and the noise transmission to be predicted so that effective treatments can be applied. Among these techniques, the finite element method and the boundary element method are at present the most important ones. However, these methods are not suited to the analysis of the behaviour of systems in the mid- and high-frequency range, because a small mesh size is required which makes model generation, turnaround time, and computations too costly.

At high frequencies where the modal density of structures is relatively high, statistical energy analysis (SEA)[1] is often used to predict the space and frequency-averaged energy level of each component of a built-up structure. However, the SEA method gives no information on the spatial repartition of energies within each substructure. In addition, SEA requires the use of coupling loss factors which are difficult to predict for complex structures.

Concerning the dynamic analysis of structures in the mid- and high-frequency range, some methods are still based upon the modal behavior. Dowell and Kubota[2] and Doherty and Dowell[3] have shown that some results of SEA can be derived from an asymptotic limit of classical modal analysis. They called this approach the asymptotic modal analysis (AMA). This technique is based on the assumptions that in the frequency domain under consideration, the modal characteristics such as masses, frequencies, and damping, do not vary rapidly. Thus at high frequencies, AMA can be used to predict frequency-averaged vibrational responses. However, the feasibility of this technique for the analysis of complex structures has yet to be demonstrated.

In an another context, Guyader[4] developed the so-called modal sampling method (MSM) for single and coupled rods. It consists of retaining only the most energetic modes in the modal decomposition of the structural response. Then, the method allows the prediction of an average level of structural response. This technique is still under development.

The power flow method presented by Belov *et al.*[5] and Luzzato[6] would appear to be an interesting alternative for use in the mid- and high-frequency range. This approach takes into account the spatial variation of energy within systems and is an analytical formulation. In addition it is amenable to

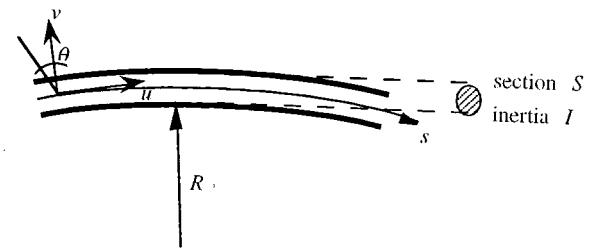


FIG. 1. Naturally curved beam.

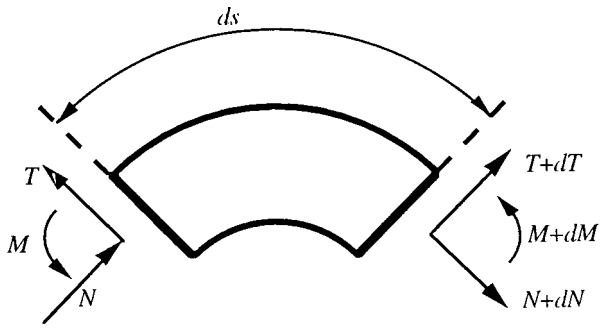


FIG. 2. Differential element from a curved beam.

a numerical implementation such as finite or boundary element methods. Nefske and Sung[7] established the relation between the energy density and its flow in beams, and wrote a second-order differential equation similar to that which describes heat conduction. Wohlever and Bernhard,[8] and Bouthier and Bernhard[9,10] investigated the procedure further. They proposed an energy model for bars, beams, membranes, plates, and acoustical cavities. The cases considered involved a single mode of propagation (unique group velocity), and a single form of energy.

The aim of this study is to present an energy model of Love–Kirchhoff curved beams.[11] First, the equation of motion is established in terms of radial and tangential displacement of a curved beam. The expressions of the total energy density and the energy flow are also provided. Energy equation and boundary conditions are developed separately for flexural and longitudinal waves. Finally, a complete energy model for curved beams including both flexural and longitudinal effects is given. A numerical example is presented to show the feasibility of this method.

I. KINEMATIC MODEL FOR CURVED BEAMS

The addition of curvature to beams which leads to an infinite variety of shapes, significantly modifies the vibrational behavior of beams. The standard reference for the governing equations of beams of arbitrary curvature is Love.[11] Referring to Love, the dynamical equation of motion for a beam of constant curvature is presented with several simplifications. Energy quantities are also given.

First of all, the curved beam is assumed to be excited by a harmonic pure tone $f=\omega/2\pi$ and steady state conditions are assumed. Therefore, by using complex notation, the time dependence $e^{i\omega t}$ is suppressed in the remaining text. A hysteretic damping model is adopted. Thus a complex modulus $E=E_0(1+i\eta)$ is introduced where E_0 is the Young's modulus.

Consider the curved beam shown in Fig. 1. The kinematic behavior depends on the displacement of each section. Three degrees of freedom are needed to describe displacements. The tangential displacement along the neutral axis s of the beam is denoted by u, and ν and θ denote the radial displacement and the rotation, respectively, of each section. The tensile load is noted N, the shear force T, and the bending moment M. The variations of these quantities are shown for a positive increment of arc length ds in Fig. 2.

The dynamic equilibrium of the element is then written in the three directions of displacement:

$$R\frac{dN}{ds}+T=-\rho SR\omega^2 u,$$

$$R\frac{dT}{ds}-N=-\rho SR\omega^2 \nu, \tag{1}$$

$$\frac{dM}{ds}+T=-\rho I\omega^2\theta.$$

It can be also shown[11] that the relationships between the tensile load N, the bending moment M, and the displacements are as follows:

$$N=ES\left(\frac{\nu}{R}+\frac{du}{ds}\right),\quad M=EI\frac{d}{ds}\left(-\frac{u}{R}+\theta\right). \tag{2}$$

In addition, assuming that the plane sections remain orthogonal to the neutral axis, the relationship which relates the degree of freedom θ to the radial displacement ν is simply

$$\theta=\frac{d\nu}{ds}. \tag{3}$$

By neglecting the inertia term of rotation in the momentum equilibrium relationship (1), and by combining the expressions (1)–(3), the governing equations of a curved beam under steady state conditions, become

$$EI\frac{d^2}{ds^2}\left(\frac{u}{R}-\frac{d\nu}{ds}\right)+ES\frac{d}{ds}\left(\nu+R\frac{du}{ds}\right)=-\rho RS\omega^2 u,$$

$$EI\frac{d^3}{ds^3}\left(\frac{u}{R}-\frac{d\nu}{ds}\right)-ES\left(\frac{\nu}{R^2}+\frac{1}{R}\frac{du}{ds}\right)=-\rho S\omega^2\nu. \tag{4}$$

For the following developments, the propagation characteristics are obtained for such structures, by considering harmonic waves of the form

$$u=ae^{-iks},\quad \nu=be^{-iks}. \tag{5}$$

By introducing these expressions into the governing equations (4), the following matrix equation is obtained:

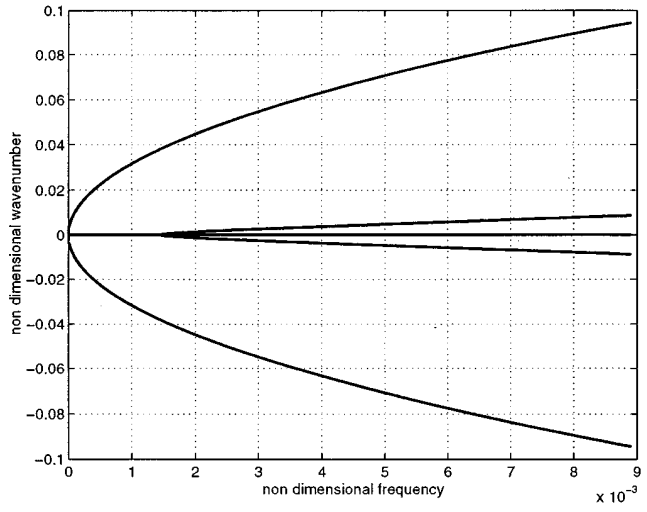


FIG. 3. Dispersion curve of a curved beam: nondimensional wave number $\overline{k}$ versus nondimensional frequency $\overline{\omega}$ for $\overline{\epsilon}=0.0014$.

$$\begin{bmatrix} \frac{\omega^2R^2}{c_0^2}-k^2R^2-k^2\mu^2 & -i(kR+k^3R\mu^2) \\ i(kR+k^3R\mu^2) & \frac{\omega^2R^2}{c_0^2}-1-k^4R^2\mu^2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \tag{6}$$

where $c_0 \overset{\text{def}}{=} \sqrt{E/\rho}$ and $\mu \overset{\text{def}}{=} \sqrt{I/S}$.

Freely propagating harmonic waves may exist only if the determinant in the above system vanishes. The resulting dispersion equation is

$$\left(\frac{\omega^2R^2}{c_0^2}-k^2R^2-k^2\mu^2\right)\left(\frac{\omega^2R^2}{c_0^2}-1-k^4R^2\mu^2\right) -(kR+k^3R\mu^2)^2=0. \tag{7}$$

This bicubic equation, where the unknown is k, leads to six complex solutions where three wave numbers have opposite signs. The explicit expressions for these complex wave numbers can be obtained by using the Cardan formulas, but they are not given here for the sake of simplicity. However, some asymptotic expressions are derived for small values of $\overline{\epsilon}=\mu/R$ in Appendix A. The real parts of the nondimensional wave numbers $\overline{k}=k\mu$ deduced from the dispersion equation are plotted in Fig. 3 versus nondimensional frequency $\overline{\omega}=\mu\omega/c_0$. This curve depends on the nondimensional parameter $\overline{\epsilon}=\mu/R$ and shows the existence of two different regions. In the first region, there are two symmetric branches of propagation, while in the second region four propagating modes are exhibited. These remarks will be very useful for the wave analysis given below. We are interested in modeling at medium- and high-frequencies in the second region.

The wave number solutions of (7) can be noted as follows:

$$k_l, -k_l, k_f^p, -k_f^p, k_f^e, -k_f^e, \tag{8}$$

where the subscript l refers to longitudinal waves and f to flexural waves. In addition, the superscript p refers to propagating waves while e refers to evanescent waves. In Appendix A, it can be seen that, when the radius of curvature R becomes very large (i.e., R tends to infinity), the solutions in Eq. (8) become

$$\omega\sqrt{\frac{\rho}{E}}, -\omega\sqrt{\frac{\rho}{E}}, \left(\frac{\rho S}{EI}\omega^2\right)^{1/4}, -\left(\frac{\rho S}{EI}\omega^2\right)^{1/4}, i\left(\frac{\rho S}{EI}\omega^2\right)^{1/4}, -i\left(\frac{\rho S}{EI}\omega^2\right)^{1/4}. \tag{9}$$

These expressions are the wave numbers for longitudinal motion of bars and transverse motion of Euler–Bernoulli straight beams. This justifies the use of the subscript l and f.

The total energy density is the sum of energies associated with the longitudinal and flexural motions, as follows:

$$W_l=\frac{\rho S\omega^2}{4}uu^*+\frac{E_0S}{4}\left(\frac{\nu}{R}+\frac{du}{ds}\right)\left(\frac{\nu}{R}+\frac{du}{ds}\right)^*, \qquad W_f=\frac{\rho S\omega^2}{4}\nu\nu^*+\frac{E_0I}{4}\frac{d}{ds}\left(\frac{u}{R}-\frac{d\nu}{ds}\right)\frac{d}{ds}\left(\frac{u}{R}-\frac{d\nu}{ds}\right)^*. \tag{10}$$

Moreover, there are three kinds of active energy flow which represent the flexural and longitudinal behavior, and the energy flow density exchanged between tangential (longitudinal) and radial (flexural) waves, respectively,

$$P_l=\text{Real}\left(\frac{i\omega}{2}\left\{N-\frac{M}{R}\right\}u^*\right), \qquad P_f=\text{Real}\left(\frac{i\omega}{2}\left\{T\nu^*+M\frac{d\nu^*}{ds}\right\}\right), \qquad p_{lf}=\text{Real}\left(\frac{i\omega}{2R}\left\{N\nu^*+M\frac{du^*}{ds}\right\}\right), \qquad p_{fl}=-p_{lf}. \tag{11}$$

In the following sections, several energy models will be presented. The aim of these models is to be able to predict the energy quantities that incorporate several simplifications without the need to solve the dynamical equations of motion (4).

II. PREAMBLE FOR THE ENERGY MODELS

We now present the smooth energy formulation (SEF) for curved beams. This approach is based on the following assumptions:

(1) The damping loss factor is small ($\eta \ll 1$).
(2) The participation of evanescent waves can be neglected.
(3) The interferences between propagating waves can be neglected.

Two energy variables are involved in these models: the energy density W defined as the sum of the kinetic energy density and the potential energy density, and the active energy flow P defined as the real part of the complex energy flow. As in previous section, a subscript is added that refers to the type of wave: l for longitudinal and f for flexural. Moreover, a superscript $+$ or $-$ denotes the direction of propagation of the wave at hand. We define the incident

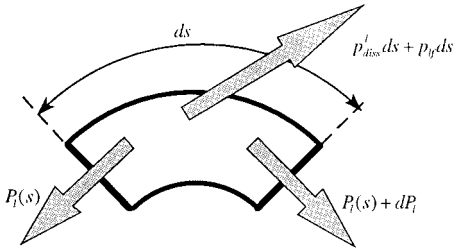


FIG. 4. Local power balance.

energy density W^+ (resp. reflected energy density W^-) and the incident energy flow P^+ (esp. reflected energy flow P^-) as the energy quantities supported by a simple incident wave (resp. reflected wave), traveling in a positive (resp. negative) direction. These quantities are sometimes called partial energies.

Obviously, when an incident wave and a reflected wave occur simultaneously, the global energy density W and the energy flow P are not *a priori* simply the sum of the corresponding partial energies. An additional term appears because of the existence of interferences between the two waves. As the smooth energy formulation (SEF) does not take into account those interferences, this additional term is neglected and the global energy density W and the active energy flow P are given in the following form:

$$P = P^+ + P^-, \quad W = W^+ + W^-. \tag{12}$$

This is a linear superposition principle extended to energy quantities.

In the following sections three energy models are studied. The first one considers just the longitudinal energies produced by a tangential wave where the effect of flexural energies is neglected. The second one considers just the flexural energies produced by a flexural wave. Finally, the third one is the most complete and takes into account the effects of the exchange of energy between the flexural and the longitudinal waves. Indeed, the previous kinematic equations show that if a longitudinal wave or a flexural wave appear alone, both longitudinal and flexural energies are produced. The two first models must not be considered as particular cases of the third but actually as an illustration. However, this approximation is introduced for the simplification that it brings. As can be seen, the power exchanged between the two waves is not high especially at high frequencies. If a given system is excited with a tangential force or with a radial force, the corresponding energy dominates and the other energy may be neglected.

III. ENERGY MODEL FOR TANGENTIAL WAVE

A. Energy flow equation

Consider a system subjected to only longitudinal vibrations. The energy quantities associated with this wave are denoted as W_l and P_l.

The first step in the derivation of the energy models is the well-known local power balance (Fig. 4). Under steady state conditions with no load, the global power balance becomes

$$\frac{dP_l}{ds} + p^l_{\text{diss}} + p_{lf} = 0 \quad \text{with} \quad p^l_{\text{diss}} = \eta\omega W_l \quad \text{and} \quad p_{lf} = a_l W_l. \tag{13}$$

The second relationship (13) is a common model for dissipation. For instance, this is precisely the model used in statistical energy analysis.[1] It assumes no distinction between kinetic and potential energy densities. While the third relationship (13) is the part of longitudinal energy transferred to flexural energy, flexural energy is not taken into account in this model, and may be considered as a dissipative term.

Hence, in order to establish an explicit energy equation, a relationship between the global active energy flow and the total energy density $P_l = H(W_l)$ has to be obtained. This relationship depends on the kind and the number of the waves that occur which, in this section, is just the longitudinal wave. It is not easy to exhibit the more general energy operator H. However, an alternative is to use the partial energy quantities. Then, the partial power balances for separate incident energies and reflected energies can be written as

$$\frac{dP_l^+}{ds} + (\eta\omega + a_l) W_l^+ = 0 \quad \text{and} \quad \frac{dP_l^-}{ds} + (\eta\omega + a_l) W_l^- = 0. \tag{14}$$

Obviously, the global power balance given in expression (13) can be deduced by summing these relationships. In addition, using the partial quantities, the relationships between the partial energy density and active energy flow are merely a proportionality. Then

$$P_l^+ = c_g^l W_l^+ \quad \text{and} \quad P_l^- = -c_g^l W_l^-, \tag{15}$$

where c_g^l is the group velocity associated to the longitudinal wave.

The group velocity is the speed at which energy flows. Its definition in terms of the longitudinal wave number is

$$c_g^l = \frac{d\omega}{dk_l}. \tag{16}$$

Alternatively, its value can be evaluated directly from the proportionality constants given in Appendix B.

The partial constitutive laws given in Eq. (15) are established for a nondissipative wave guide[12] since the employed group velocity is that for an undamped system. For SEF developments, the damping loss factor is assumed to be very small (first assumption), and then relationships (15) are assumed to remain valid. In addition, the effects of damping are taken into account by way of the power being dissipated in the power balance relations (14).

Substracting the relationships (14) yields

$$\frac{d}{ds}(P_l^+ - P_l^-) + (\eta\omega + a_l)(W_l^+ - W_l^-) = 0.$$

Using proportionality relationships (15)

$$c_g^l \frac{d}{ds}(W_l^+ + W_l^-) + \frac{\eta\omega + a_l}{c_g^l}(P_l^+ + P_l^-) = 0.$$

Thus by virtue of (12), the global active energy flow, P_l, is simply proportional to the first derivative of the global energy density W_l:

$$P_l = -\frac{c_g^{l^2}}{\eta\omega + a_l}\frac{dW_l}{ds}. \tag{17}$$

Finally, combining expressions (17) and (13), one obtains the explicit energy equation for tangential movement of curved waveguide:

$$\frac{d^2W_l}{ds^2} - \frac{(\eta\omega + a_l)^2}{c_g^{l^2}} W_l = 0. \tag{18}$$

Let us define the real wave number k_∞^l as the wave number of the straight longitudinal excited nondissipative system: $k_\infty^l \overset{\text{def}}{=} \omega\sqrt{\rho/E_0}$. Then, when the nondimensional quantity Rk_∞^l becomes infinite $Rk_\infty^l \to \infty$, the group velocity c_g^l is the same as for straight bar, ω/k_∞^l, and the exchanged coefficient a_l vanishes, so energy equation (18) is the same as that developed in Refs. 8 and 13:

$$\frac{d^2W}{ds^2} - \eta^2 k_\infty^l W = 0. \tag{19}$$

The solution of Eq. (18) is the sum of an incident and a reflected term:

$$W_l(s) = A^+ e^{[(\eta\omega + a_l)/c_g^l]s} + A^- e^{[(\eta\omega + a_l)/c_g^l]s}. \tag{20}$$

To solve this problem completely, one must determine the energy amplitudes A^+ and A^- occurring in the previous expression (20). To calculate the energy magnitudes A^+ and A^-, the classical boundary conditions may be written in terms of the variables W_l and P_l. A complete study concerning the energetic boundary and coupling conditions has been already made[14] in the case of a simple Euler–Bernoulli beam and other waveguides.[15] The general procedure proposed there remains, of course, valid in the case of radial waves in the waveguide. For instance, at an excited node located at s_0, the active energy flow is assumed to be known.

$$P_l(s_0) = P_{inj}. \tag{21}$$

For a dissipative end located at s_1, the equivalent energy boundary condition is

$$P_l^-(s_1) = r_l P_l^+(s_1)$$

or

$$(P_l - c_g^l W_l)(s_0) = r_l (P_l + c_g^l W_l)(s_0), \tag{22}$$

where r_l is a reflection coefficient. Note that this coefficient may be smaller than one even for a nondissipative end due to mode conversion: A part of longitudinal incident energy flow is converted into flexural energy flow.

Boundary conditions (21) and (22), when combining with relationship (20), allow the energy density magnitude A^+ and A^- to be calculated.

B. The envelope energy expression

Energy equation (18) given above allows the prediction of the level of the total energy density, and the active energy flow without interferences. Its solution (20) has a smooth spatial variation with no local oscillations. In this section, further information is deduced from the energy equation (18) by comparing it to the solution of the equations of motion (4).

Indeed, the particular solution of Eqs. (4), corresponding to a couple of longitudinal waves, can be written

$$\begin{aligned} u &= a^+ e^{-ik_l s} + a^- e^{ik_l s}, \\ \nu &= b^+ e^{-ik_l s} + b^- e^{ik_l s}. \end{aligned} \tag{23}$$

By virtue of (6), a proportionality exists between $b^\pm$ and $a^\pm$. Let ξ_l denote this proportionality, so that $b^+ = \xi_l a^+$ and $b^- = -\xi_l a^-$ where

$$\xi_l = \frac{\omega^2 R^2/c_0^2 - k_l^2 R^2 - k_l^2\mu^2}{i(k_l R + k_l^3 R\mu^2)} = \frac{-i(k_l R + k_l^3 R\mu^2)}{\omega^2 R^2/c_0^2 - 1 - k_l^4 R^2\mu^2}. \tag{24}$$

By introducing expression (23) in the longitudinal energy density expression given by (10), the following formulas for the total energy density are obtained:

$$\begin{aligned} W_l(s) &= W_{\text{smooth}}(s) + W_{\text{asc}}(s), \\ W_{\text{smooth}}(s) &= \frac{1}{4}\left[\rho S\omega^2 + E_0 S\left|\frac{\xi_l}{R} - ik_l\right|^2\right] \\ &\quad \times (|a^+|^2 e^{2\,\text{Imag}(k_l)s} + |a^-|^2 e^{-2\,\text{Imag}(k_l)s}), \\ W_{\text{asc}}(s) &= \frac{1}{4}\left[\rho S\omega^2 - E_0 S\left|\frac{\xi_l}{R} - ik_l\right|^2\right] \\ &\quad \times (a^+ a^{-*} e^{-2i\,\text{Real}(k_l)s} + a^- a^{+*} e^{2i\,\text{Real}(k_l)s}). \end{aligned} \tag{25}$$

The comparison between the total energy density (25) calculated using solution (23) and the one given in expression (20) and evaluated from the smooth energy formulation, shows that relationship (20) is the smooth term that occurs in (25). This first term, corresponding to the real exponential terms, has a slow spatial variation, while the second energy term has an oscillating spatial variation. In addition, a mere identification gives the magnitudes A^+ and A^- in terms of $|a^+|^2$ and $|a^-|^2$:

$$\begin{aligned} A^+ &= \frac{1}{4}\left[\rho S\omega^2 + E_0 S\left|\frac{\xi_l}{R} - ik_l\right|^2\right]|a^+|^2, \\ A^- &= \frac{1}{4}\left[\rho S\omega^2 + E_0 S\left|\frac{\xi_l}{R} - ik_l\right|^2\right]|a^-|^2. \end{aligned} \tag{26}$$

Moreover, the magnitude of the oscillating components W_{osc} is given by

$$\begin{aligned} M_{\text{asc}} &= \frac{1}{2}\left|\rho S\omega^2 - E_0 S\left|\frac{\xi_l}{R} - ik_l\right|^2\right||a^- a^{+*}| \\ &= 2\frac{\left|\rho S\omega^2 - E_0 S\left|\frac{\xi_l}{R} - ik_l\right|^2\right|}{\left|\rho S\omega^2 + E_0 S\left|\frac{\xi_l}{R} - ik_l\right|^2\right|}\sqrt{A^- A^+}, \end{aligned} \tag{27}$$

where A^+ and A^- are the energy magnitudes predicted by SEF.

In the case of large value of Rk_∞^l, the magnitude of the oscillating components is simply

$$M_{\text{asc}} \approx 2\sqrt{AB}/R^2 k_\infty^{l^2}. \tag{28}$$

As a result of this remark, the energy equation (20) contains additional information on the lower and the upper energy envelopes:

$$W_{\text{upper}}(s)=A^{+}e^{[(\eta\omega+a_l)/c_g^l]s}+A^{-}e^{[(\eta\omega+a_1)/c_g^l]s}+M_{\text{asc}},$$

$$W_{\text{lower}}(s)=A^{+}e^{[(\eta\omega+a_l)/c_g^l]s}+A^{-}e^{[(\eta\omega+a_l)/c_g^l]s}-M_{\text{asc}}. \tag{29}$$

Consequently, SEF for tangential waves predicts not only the smooth level of the total energy density and active energy flow, but their upper and lower envelopes too.

IV. ENERGY MODEL FOR RADIAL WAVES

In this section an energy model for a system containing only radial waves is developed in a similar way as for the system with only tangential waves in Sec. III.

A. Energy equation

Let us consider the solution of the frequency equation. In Eq. (8) we noted that two kinds of wave numbers exist for radial case: an evanescent wave number k_f^e and a propagating wave number k_f^p. The amplitudes of the evanescent waves decrease rapidly away from boundaries, and are then neglected for SEF developments (second assumption). Therefore the only group velocity to be considered for radial waves is associated to the wave number for propagating waves:

$$c_g^f=\frac{d\omega}{dk_f^p}. \tag{30}$$

Hence, the problem associated with the propagating radial waves becomes similar to the tangential case treated before. In particular, the relationship between the global active energy flow and total energy density is as follows:

$$P_f=-\frac{c_g^{f^2}}{\eta\omega+a_f}\frac{dW_f}{ds}. \tag{31}$$

By introducing this constitutive law in the active power balance, one obtains the following energy equation for the radial behavior of a curved waveguide:

$$\frac{d^2W_f}{ds^2}-\frac{(\eta\omega+a_f)^2}{c_g^{f^2}}W_f=0. \tag{32}$$

By introducing the wave number k_∞^f corresponding to a straight nondissipative Euler–Bernoulli beam, $k_\infty^f \underset{\text{def}}{=}(\rho S\omega^2/E_0I)^{1/4}$, and by assuming the quantity Rk_∞^f to be very large, the energy equation obtained is merely the well-known energy equation developed by Refs. 7, 8, and 13:

$$\frac{d^2W}{ds^2}-\frac{\eta^2k_\infty^{f^2}}{4}W=0. \tag{33}$$

The solution of energy equation (32) is

$$W_f(s)=A^{+}e^{-[(\eta\omega+a)/c_g^f]s}+A^{-[(\eta\omega+a)/c_g^f]s}. \tag{34}$$

The equation above is solved using energetic boundary conditions. Here, relationships similar to (21) and (22) given in the tangential case are valid.

B. The energy envelope expression

In a manner similar to that used for tangential waves, the upper and lower total energy density envelopes can be determined. Let us consider a couple of propagating flexural waves where the resulting tangential and radial displacements are as follows:

$$u=a^{+}e^{-ik_f^ps}+a^{-}e^{ik_f^ps},$$
$$\nu=b^{+}e^{-ik_f^ps}+b^{-}e^{ik_f^ps}. \tag{35}$$

The proportionality between $b^{\pm}$ and $a^{\pm}$ still applies i.e., $b^{+}=\xi_f a^{+}$ and $b^{-}=-\xi_f a^{-}$ where

$$\xi_f=\frac{\omega^2R^2/c_0^2-k_f^{p^2}R^2-k_f^{p^2}\mu^2}{i(k_f^pR+k_f^{p^3}R\mu^2)}$$
$$=\frac{-i(k_f^pR+k_f^{p^3}R\mu^2)}{\omega^2R^2/c_0^2-1-k_f^{p^4}R^2\mu^2}. \tag{36}$$

Then, the expression for the transverse energy density in case of a flexural propagating wave is:

$$W_f(s)=W_{\text{smooth}}(s)+W_{\text{asc}}(s),$$
$$W_{\text{smooth}}(s)=\frac{1}{4}\left[\rho S\omega^2|\xi_f|^2+E_0I\left|\frac{ik_f^p}{R}-k_f^{p^2}\xi_f\right|^2\right]$$
$$\times(|a^{+}|^2e^{2\,\text{Imag}(k_f^p)s}+|a^{-}|^2e^{-2\,\text{Imag}(k_f^p)s}),$$
$$W_{\text{asc}}(s)=\frac{-1}{4}\left[\rho S\omega^2|\xi_f|^2+E_0I\left|\frac{ik_f^p}{R}-k_f^{p^2}\xi_f\right|^2\right]$$
$$\times(a^{+}a^{-*}e^{-2i\,\text{Real}(k_f^p)s}+a^{-}a^{+*}e^{2\,\text{Real}(k_f^p)s}). \tag{37}$$

Hence, as in the tangential case, the magnitude of the oscillating field is given by

$$M_{\text{asc}}=2\sqrt{A^{+}A^{-}}. \tag{38}$$

It is surprising that this magnitude does not depend explicitly on the presence of curvature. In fact, the obtained magnitude of oscillations $2\sqrt{A^{+}A^{-}}$ is exactly the one given in Ref. 16 for the case of a straight Euler–Bernoulli beam. Obviously, the curvature affects the values of the coefficients A^{+} and A^{-} by means of the group velocity which depends on this curvature.

In the case of radial waves in curved wave guides, the upper and lower total energy density variation are given simply by

$$W_{\text{upper}}(s)=A^{+}e^{[(\eta\omega+a_f)/c_g^f]s}+A^{-}e^{[(\eta\omega+a_f)/c_g^f]s}+M_{\text{asc}},$$

$$W_{\text{lower}}(s)=A^{+}e^{[(\eta\omega+a_f)/c_g^f]s}+A^{-}e^{[(\eta\omega+a_f)/c_g^f]s}-M_{\text{asc}}. \tag{39}$$

V. COMPLETE ENERGY MODEL OF CURVED BEAMS

The case of the simultaneous propagation of radial and tangential waves in the waveguide is more complicated than the propagation of the individual waves. Every kind of energy quantity which was previously defined in Sec. I, appears simultaneously in this case:

$$P_l^+, P_l^-, P_f^+, P_f^-, W_l^+, W_l^-, W_f^+, W_f^- . \tag{40}$$

To these quantities, we add the active energy flow, which characterizes the exchange between the longitudinal (tangential waves) and the flexural (radial waves) behavior. The exchanged energy flows are

$$p_{lf}^+, p_{lf}^- . \tag{41}$$

Let us generalize the derivation of the energy equations established for a single wave. We noted that when interferences were neglected, a linear superposition principle is valid for energy quantities. Relationship (12) then becomes

$$Z = Z^+ + Z^- \quad \text{with} \quad Z = P, \ W, \ \text{or} \ p. \tag{42}$$

In this case, the local power balance is composed of two coupled equations due to the presence of the exchanged energy flows:

$$\begin{aligned} &\frac{d}{ds}\begin{pmatrix} P_l^+ \\ P_f^+ \end{pmatrix} + \begin{pmatrix} p_{\text{diss}}^{l+} \\ p_{\text{diss}}^{f+} \end{pmatrix} + \begin{pmatrix} p_{lf}^+ \\ -p_{lf}^+ \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \\ &\frac{d}{ds}\begin{pmatrix} P_l^- \\ P_f^- \end{pmatrix} + \begin{pmatrix} p_{\text{diss}}^{l-} \\ p_{\text{diss}}^{f-} \end{pmatrix} + \begin{pmatrix} p_{lf}^- \\ -p_{lf}^- \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \end{aligned} \tag{43}$$

The damping model, $p_{\text{diss}} = \eta\omega W$ is still valid. Now, the constitutive relationships between the propagating quantities have to be established. In the coupled case, relationships (15) used for both radial waves and tangential waves are still valid, but the proportionality coefficient becomes a 2 by 2 matrix:

$$\begin{pmatrix} P_l^+ \\ P_f^+ \end{pmatrix} = [C_g]_{2\times 2}\begin{pmatrix} W_l^+ \\ W_f^+ \end{pmatrix}, \quad \begin{pmatrix} P_l^- \\ P_f^- \end{pmatrix} = -[C_g]_{2\times 2}\begin{pmatrix} W_l^- \\ W_f^- \end{pmatrix}. \tag{44}$$

An additional relationship is written for the exchanged energy flows as follows:

$$\begin{pmatrix} p_{lf}^+ \\ -p_{lf}^+ \end{pmatrix} = [A]_{2\times 2}\begin{pmatrix} W_l^+ \\ W_f^+ \end{pmatrix}, \quad \begin{pmatrix} p_{lf}^- \\ -p_{lf}^- \end{pmatrix} = [A]_{2\times 2}\begin{pmatrix} W_l^- \\ W_f^- \end{pmatrix}, \tag{45}$$

where $[A]$ is a 2 by 2 matrix with a second row equal to the opposite of the first:

$$[A] = \begin{pmatrix} \lambda & \gamma \\ -\lambda & -\gamma \end{pmatrix}.$$

Note that the matrix appearing in the second relationship (45) is not the opposite of those in the first relationship contrary to Eq. (44). This means that when the propagation direction of a wave is inverted, there is no modification of the active energy flow exchanged between the radial and tangential waves. The complete expressions of the matrices $[C_g]$ and $[A]$ are given in the Appendix B.

Given the constitutive relationships in Eqs. (44) and (45), it becomes easy to deduce the equation for the energy flow for coupled radial and tangential waves. The sum of expressions (43) leads to the global power balance:

$$\frac{d}{ds}\begin{pmatrix} P_l \\ P_f \end{pmatrix} + \eta\omega\begin{pmatrix} W_l \\ W_f \end{pmatrix} + [A]_{2\times 2}\begin{pmatrix} W_l \\ W_f \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{46}$$

The difference between the expressions yields the constitutive relation between the global energetic variables:

$$\begin{aligned} &\frac{d}{ds}\left\{\begin{pmatrix} P_l^+ \\ P_f^+ \end{pmatrix} - \begin{pmatrix} P_l^- \\ P_f^- \end{pmatrix}\right\} + (\eta\omega[I]_{2\times 2} + [A]_{2\times 2}) \\ &\times\left\{\begin{pmatrix} W_l^+ \\ W_f^+ \end{pmatrix} - \begin{pmatrix} W_l^- \\ W_f^- \end{pmatrix}\right\} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \end{aligned} \tag{47}$$

where $[I]$ is the identity matrix. Substituting (44) for each term yields

$$\begin{pmatrix} P_l \\ P_f \end{pmatrix} = -[C_g](\eta\omega[I] + [A])^{(-1)}[C_g]\frac{d}{ds}\begin{pmatrix} W_l \\ W_f \end{pmatrix}. \tag{48}$$

Thus by substituting expressions (48) into (46), the system of energy equations for the coupled problem is written as

$$\frac{d^2}{ds^2}\begin{pmatrix} W_l \\ W_f \end{pmatrix} - ([C_g]^{(-1)}(\eta\omega[I] + [A]))^2\begin{pmatrix} W_l \\ W_f \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{49}$$

The solutions of this system introduce four energy magnitudes which may be evaluated from energy boundary conditions. For instance, at an excited node, the injected active energy flow is assumed to be known, and then

$$\begin{pmatrix} P_l \\ P_f \end{pmatrix}(s_0) = \begin{pmatrix} P_{l_{\text{inj}}} \\ P_{f_{\text{inj}}} \end{pmatrix}. \tag{50}$$

Moreover, at a dissipative end, the following relationship is written:

$$\begin{pmatrix} P_l^- \\ P_f^- \end{pmatrix}(s_1) = [R]_{2\times 2}\begin{pmatrix} P_l^+ \\ P_f^+ \end{pmatrix}(s_1)$$

or (51)

$$\begin{aligned} &\left[\begin{pmatrix} P_l \\ P_f \end{pmatrix} - [C_g]_{2\times 2}\begin{pmatrix} W_l \\ W_f \end{pmatrix}\right](s_1) \\ &= [R]_{2\times 2}\left[\begin{pmatrix} P_l \\ P_f \end{pmatrix} + [C_g]_{2\times 2}\begin{pmatrix} W_l \\ W_f \end{pmatrix}\right](s_1). \end{aligned}$$

The matrix $[R]$ is constituted from four reflection coefficients which illustrate mode-conversion phenomena. Each coefficient has a non-negative value smaller than one. The nondissipative condition leads to special relationships: The sum of each column must be equal to one.

VI. NUMERICAL SIMULATIONS

Consider a curved beam with mass density $\rho = 7800\ \text{kg/m}^3$, Young modulus $E = 2.1\ e^{11}\ \text{N/m}^2$, inertia $I = 8.3\ e^{-10}\ \text{m}^4$, cross section $S = 1\ e^{-4}\ \text{m}^2$, length $L = 1\ \text{m}$, and radius of curvature $R = 0.25\ \text{m}$. The damping ratio η is taken to be equal to 0.05.

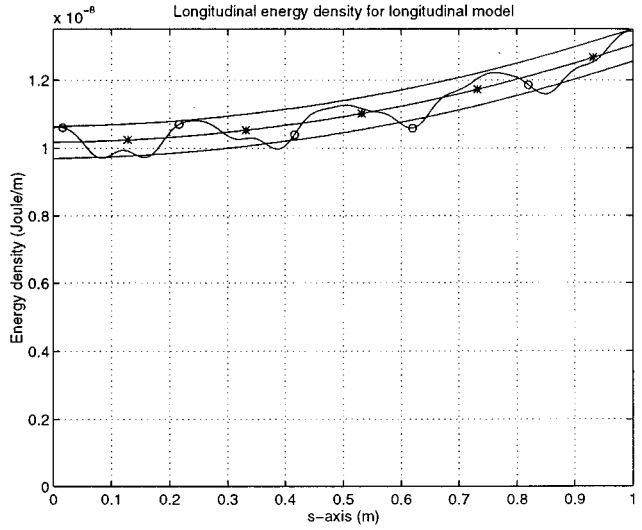


FIG. 5. Longitudinal energy density variation versus s axis; ○ equation of motion (4); * energetic equation (18) and its associated upper and lower bounds given in (29).

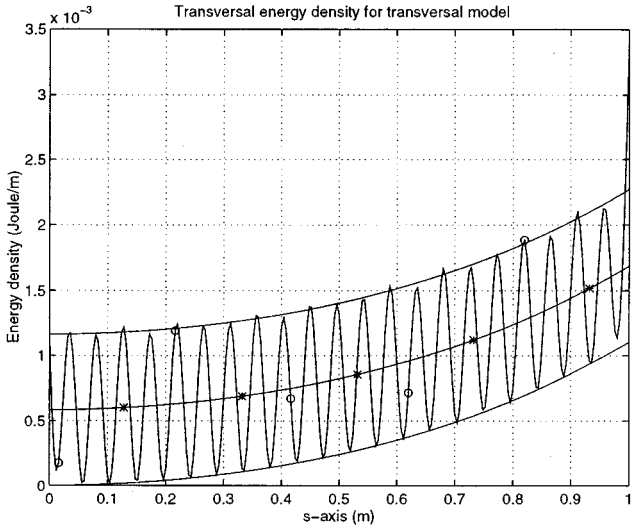


FIG. 7. Transverse energy density variation versus s axis; ○ equation of motion (4); * energetic solution (34) and its associated upper and lower bounds given in (39).

First of all, the curved beam is excited at its right end node 1 by a longitudinal force of strength 1 N. Following (21), the longitudinal active energy flow P_l is assumed to be known at node 1. At the left end of the waveguide, node 0, the clamped end boundary condition is used which corresponds to the relationship (22) with an appropriate reflection coefficient (see Appendix C). Two calculations are performed. The first solves governing equations (4) and the longitudinal energy density and energy flow are then evaluated using Eqs. (10) and (11). This classical calculation is taken as a reference. On the other hand, combining expressions (21) and (22), one can find the energy magnitudes A^+ and A^- and then the energy quantities are obtained from Eqs. (17) and (20). This is the SEF result. The results of the simulations are given in Figs. 5 and 6 for a pure-tone excitation at $f=11\,000$ Hz. Thus it can be shown that the energy model (20) produces a prediction of the local space average of total energy density and active energy flow. Therefore, as the solution given by this model varies slowly in space, a numerical implementation of this procedure with finite elements for instance, would require a few degrees of freedom. This is very attractive in the mid- and high-frequency domain. Finally, note that the energy flow predicted by SEF does not vanish at the clamped end. Actually, the local energy flow predicted by the equation of motion will disappear but its mean value over a vicinity will not disappear. This apparent dissipation is relevant to mode conversion phenomena.

Second, the curved beam is excited at node 1 by a transverse force strength 1 N. The injected active energy flow is assumed to be known at this node. At the end of this waveguide the energetic clamped end (22) is still valid at this stage. The energy system to be solved in order to determine the energy magnitudes A^+ and A^- is formally the same as in the longitudinal case. Results are given in Figs. 7 and 8. The

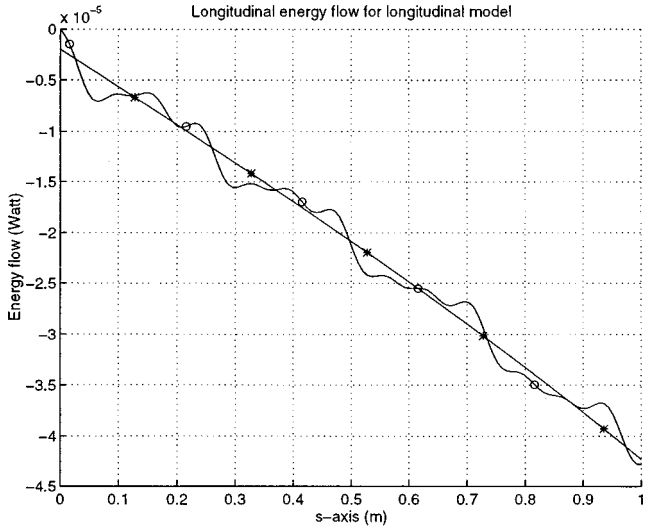


FIG. 6. Longitudinal energy flow variation versus s axis; ○ equation of motion (4); * energetic equation (18).

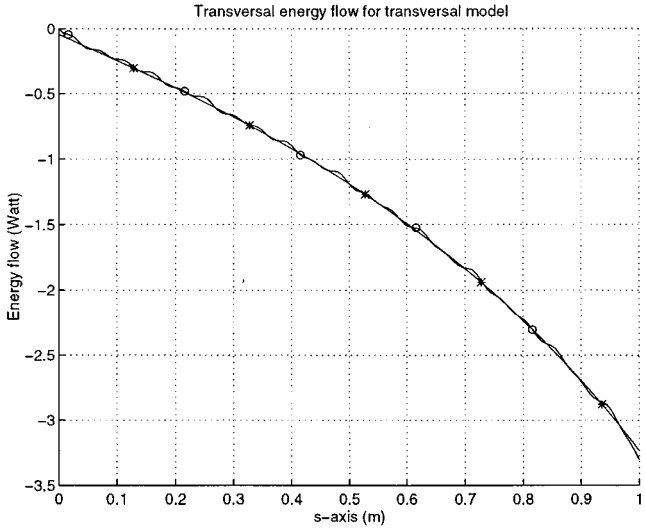


FIG. 8. Transverse energy flow variation versus s axis; ○ equation of motion (4); * energetic solution (34).

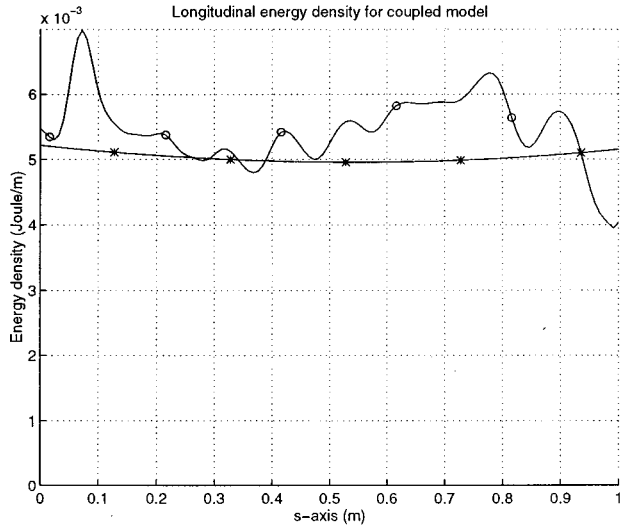


FIG. 9. Longitudinal energy density variation versus s axis; ○ equation of motion (4); * energetic equation (49).

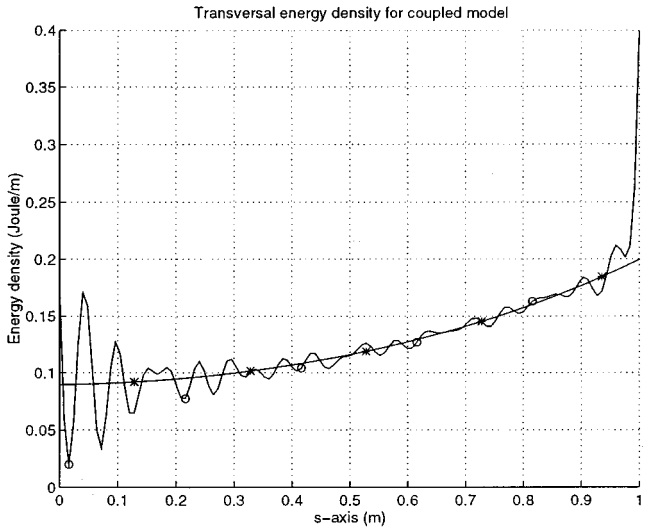


FIG. 11. Transverse energy density variation versus s axis; ○ equation of motion (4); * energetic equation (49).

comparison between the classic calculation of the energy density and the energy flow from the kinematic model (4) shows that SEF again allows the prediction of the required smooth level of the energy density and the energy flow. Moreover, the application of relationships (39) permits the upper and lower energy envelopes to be known.

Finally, the complete coupled system is simulated. Two forces are applied: a longitudinal force of strength 5 N and an antiphase transverse force of strength 1 N. The results of a simulation are given in Figs. 9–12 versus the curvilinear coordinate s. However, in opposition to previous simulations where the excitation was in pure tone, the classical calculation is now performed over the octave 5000 Hz–10 000 Hz. Actually, the matrix $[R]$ has been determined for a semi-infinite system and then it is applicable to a finite system in a frequency-average sense. However, it turns out that the differences between the pure-tone reflection coefficients (not determined here) and the semi-infinite ones are small. Then, as long as we are interested in just one kind of energy induced by respective load, a pure-tone simulation is satisfactory. But, for the coupled model, due to the large domination of longitudinal energy level by transverse energy level, a small miscalculation of the transverse reflection coefficient may induce a large difference on longitudinal energy level. This is why frequency averages are needed. Results given in Figs. 9–12 are the average over this octave. The energetic calculation stems from (49) with boundary conditions (50) and (51). However, the injected power appearing in (50) is not the exact pure-tone value at the mean frequency but the mean value for this octave. The latter is obtained by considering the infinite impedance: The injected power is the one created by the same excitation applied to the equivalent infinite curved beam. Note that the energy flows do not vanish at the clamped end as previously observed. The longitudinal energy flow at the clamped end is exactly the opposite of the

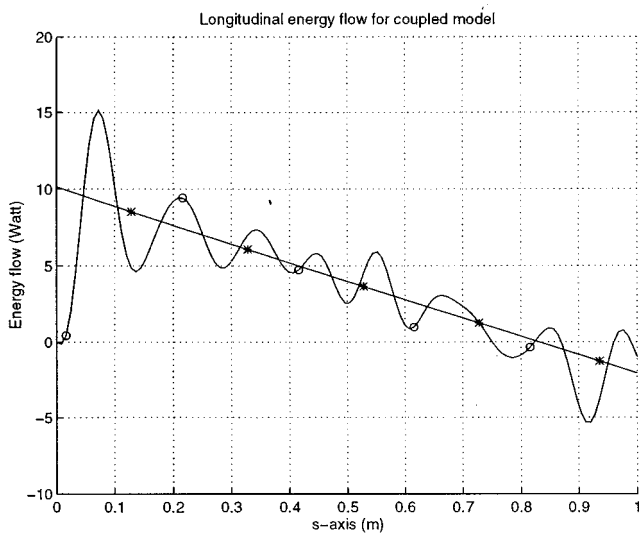


FIG. 10. Longitudinal energy flow variation versus s axis; ○ equation of motion (4); * energetic equation (49).

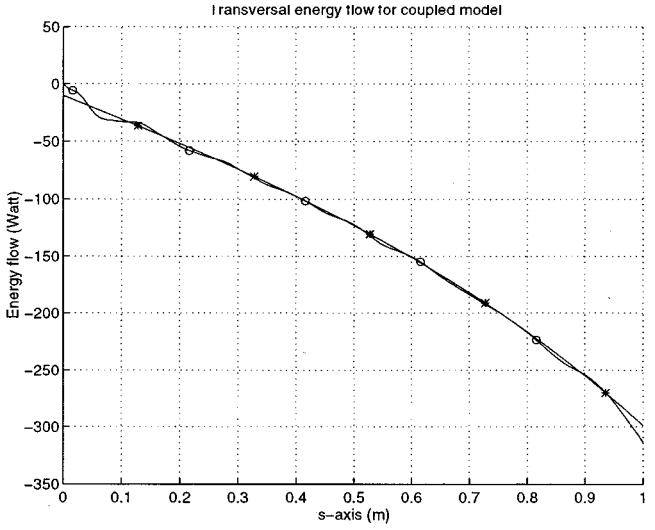


FIG. 12. Transverse energy flow variation versus s axis; ○ equation of motion (4); * energetic equation (49).

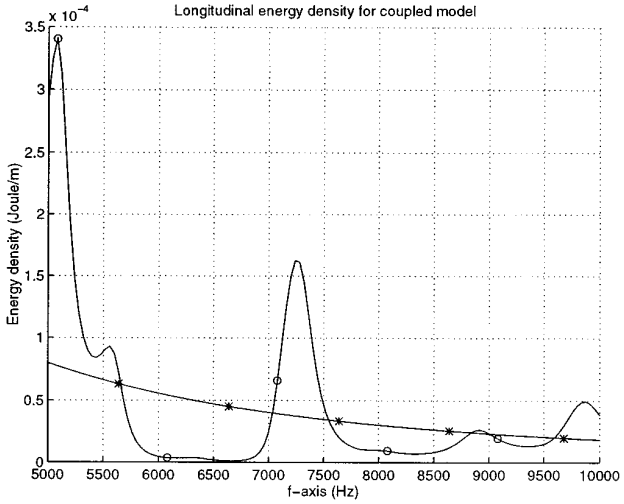


FIG. 13. Longitudinal energy density variation versus f axis; ○ equation of motion (4); * energetic equation (49).

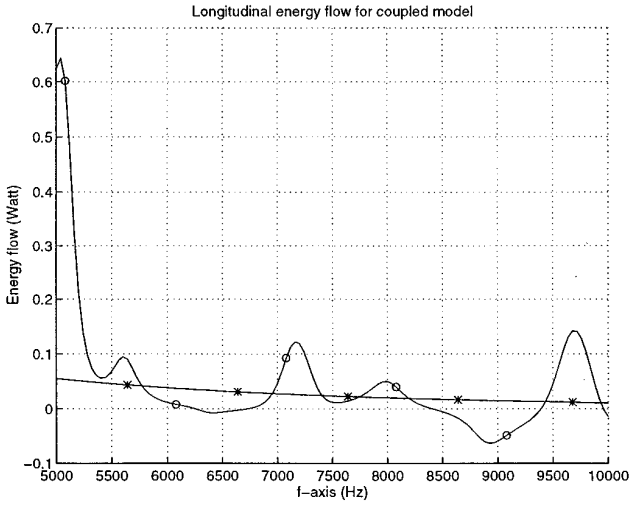


FIG. 14. Longitudinal energy flow variation versus f axis; ○ equation of motion (4); * energetic equation (49).

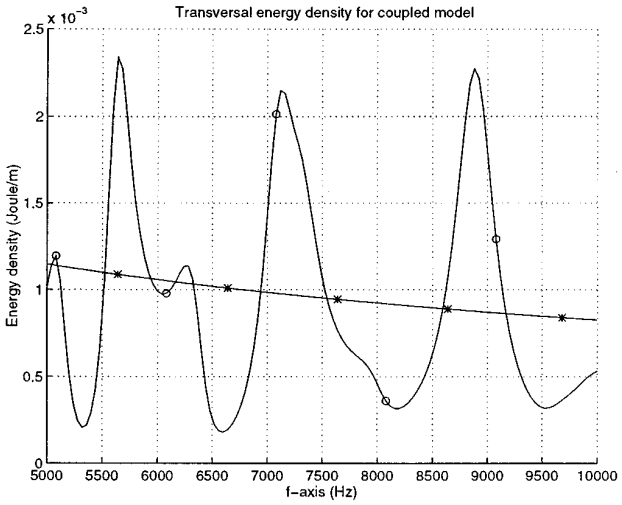


FIG. 15. Transverse energy density variation versus f axis; ○ equation of motion (4); * energetic equation (49).

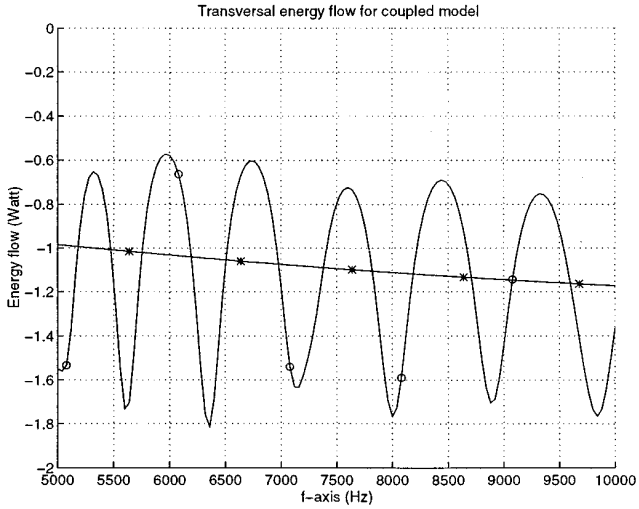


FIG. 16. Transverse energy flow variation versus f axis; ○ equation of motion (4); * energetic equation (49).

transverse energy flow at the same point. So, the clamped end is a nondissipative end. In Figs. 13–16 the energy quantities are plotted at a point versus frequency. It can be observed that the energy quantities issued from SEF are not only the local space average of exact quantities but also the frequency average. This average must be performed over a frequency band which contains at least several eigenfrequencies.

VII. CONCLUDING REMARKS

In this paper, a new formulation has been proposed to allow the prediction of the dynamical behaviour of curved beams at high frequencies. This formulation is based on a solely energetic variables which are the total energy density and the active energy flow. The results given here generalize those given in the case of straight rods and beams.[7,8,13] This method is numerically attractive in the mid- and high-frequency range.

Three models have been studied. The first one is concerned only with the longitudinal behavior. Then by considering a longitudinal wave alone, this model allows the prediction of the smooth part of the longitudinal energy density and energy flow. A similar model has been developed for the flexural behavior. However an additional assumption is required: The evanescent wave has to be neglected. In these two models, the upper and lower envelops of the energy density have been also evaluated. Finally, a complete model including the effect of the exchanged energy flow between longitudinal and flexural behavior, has been proposed.

ACKNOWLEDGMENTS

The authors gratefully acknowledge E. Luzzato (EDF–DER Département d'Acoustique et de Mécanique Vibratoire), and R. Aquilina (Centre d'Étude et de Recherche de la Discrétion Acoustique des Navires) for their kind help and advices, and R. Lambrech for editing this manuscript.

APPENDIX A

Let us rewrite frequency equation (7) by using the nondimensional quantities $\overline{k}=k\mu$, $\overline{\omega}=\mu\omega/c_0$, and $\overline{\epsilon}=\mu/R$:

$$\overline{k}^6-(\overline{\omega}^2+2\overline{\epsilon}^2)\overline{k}^4+(\overline{\epsilon}^4-\overline{\omega}^2(1+\overline{\epsilon}^2))\overline{k}^2+\overline{\omega}^2(\overline{\omega}^2-\overline{\epsilon}^2)=0. \tag{A1}$$

This is a bicubic equation for the unknown $\overline{k}$. As the wave number for a straight rod is $k_\infty^l=\omega/c_0$, the solution $\overline{k}^l$ of Eq. (A1) associated with longitudinal waves must tend to $\overline{\omega}$ as the small parameter $\overline{\epsilon}$ tends to zero. Moreover the value of $\overline{k}^l$ cannot depend on the sign of the radius of curvature R. Then, an asymptotic expression of $\overline{k}^l$ must take the form

$$\overline{k}^l=\overline{\omega}+a\overline{\epsilon}^2+o(\overline{\epsilon}^2). \tag{A2}$$

Substituting (A2) in (A1) yields

$$\overline{k}^l=\overline{\omega}+\frac{1+3\overline{\omega}^2}{2\overline{\omega}(\overline{\omega}^2-1)}\overline{\epsilon}^2+o(\overline{\epsilon}^2). \tag{A3}$$

An analogous calculation for the transverse wave number leads to

$$\overline{k}_f^p=\sqrt{\overline{\omega}}+\frac{3+\overline{\omega}}{4\sqrt{\overline{\omega}}(1-\overline{\omega})}\overline{\epsilon}^2+o(\overline{\epsilon}^2). \tag{A4}$$

These relationships are valid in case of large radius of curvature.

APPENDIX B

Let us consider a harmonic wave of form (5). The admissible values for the parameter k are given by dispersion equation (7). In order to simplify the following relationships, we note k_1 (resp. k_2), the incident longitudinal wave number k_1 (resp. the incident propagating flexural wave number k_f^p).

Moreover, referring to (6), for each harmonic wave of form (5), the corresponding displacement magnitudes a and b are proportional. So, let ξ_i designate the ratio of the displacement amplitudes b_i/a_i for $i=1,2$. The expressions for ξ_i are respectively given in Eqs. (36) and (24).

Now, let us assume that an incident radial wave and an incident tangential wave occur simultaneously in a given system. The total displacements are

$$u=a_1e^{-ik_1s}+a_2e^{-ik_2s} \quad \text{and} \quad \nu=b_1e^{-ik_1s}+b_2e^{-ik_2s}. \tag{B1}$$

Then substituting (B1) into energy expressions (10) and (11), and neglecting the cross product terms which are due to interferences between radial and tangential waves, leads to

$$W_l^+=\sum_{i=1}^{2}\frac{1}{4}\left\{\rho S\omega^2+\frac{E_0S}{R^2}|\xi_i-ik_iR|^2\right\}\times[|a_i|^2e^{2\,\mathrm{Imag}(k_i)s}],$$

$$W_f^+=\sum_{i=1}^{2}\frac{1}{4}\left\{\rho S\omega^2|\xi_i|^2+\frac{E_0I}{R^2}|k_i|^2|1+ik_i\xi_iR|^2\right\}\times[|a_i|^2e^{2\,\mathrm{Imag}(k_i)s}],$$

$$P_l^+=\sum_{i=1}^{2}\mathrm{Real}\left\{\frac{i\omega}{2}\left[\frac{ES}{R}(\xi_i-ik_iR)-\frac{EI}{R^2}ik_i(1+ik_i\xi_iR)\right]\right\}[|a_i|^2e^{2\,\mathrm{Imag}(k_i)s}], \tag{B2}$$

$$P_f^+=\sum_{i=1}^{2}\mathrm{Real}\left\{\frac{i\omega}{2}\frac{EI}{R}[-k_i^2(1+ik_i\xi_iR)\xi_i^*-|k_i|^2(1+ik_i\xi_iR)\xi_i^*]\right\}[|a_i|^2e^{2\,\mathrm{Imag}(k_i)s}],$$

$$p_{lf}^+=\sum_{i=1}^{2}\mathrm{Real}\left\{\frac{i\omega}{2}\frac{1}{R^2}[ES(\xi_i-ik_iR)\xi_i^*-EI|k_i|^2(1+ik_i\xi_iR)]\right\}[|a_i|^2e^{2\,\mathrm{Imag}(k_i)s}].$$

The corresponding quantities with a minus superscript are obtained by substituting $-k_i$ and $-\xi_i$ everywhere. Relationships (B2) can be written in a simple matrix form as

$$\begin{pmatrix}W_l^+\\W_f^+\end{pmatrix}=[w]_{2\times2}\begin{pmatrix}|a_1|^2e^{2\,\mathrm{Imag}(k_1)s}\\|a_2|^2e^{2\,\mathrm{Imag}(k_2)s}\end{pmatrix}\begin{pmatrix}P_l^+\\P_f^+\end{pmatrix}=[\wp]_{2\times2}\begin{pmatrix}|a_1|^2e^{2\,\mathrm{Imag}(k_1)s}\\|a_2|^2e^{2\,\mathrm{Imag}(k_2)s}\end{pmatrix}p_{lf}^+=[p]_{1\times2}\begin{pmatrix}|a_1|^2e^{2\,\mathrm{Imag}(k_1)s}\\|a_2|^2e^{2\,\mathrm{Imag}(k_2)s}\end{pmatrix}. \tag{B3}$$

The coefficients of the matrix $[w]_{2\times2}$, $[\wp]_{2\times2}$, and $[p]_{1\times2}$ are easily identified from expressions (B2). Thus the matrix $[C_g]_{2\times2}$ and $[A]_{2\times2}$ are given by

$$[C_g]=[\wp][w]^{(-1)} \quad \text{and} \quad [\lambda\ \gamma]=[p][w]^{(-1)}. \tag{B4}$$

When either the longitudinal or transverse model is considered, the group velocity and the coefficients a_l and a_f are determined from these matrices:

$$c_g^l=C_{g_{1,1}} \quad \text{and} \quad c_g^f=C_{g_{2,2}}, \qquad a_l=\lambda \quad \text{and} \quad a_f=-\mu. \tag{B5}$$

APPENDIX C

Let us determine the reflection coefficients matrix occurring in (51). For this purpose, an incident longitudinal wave of form (5) whose magnitudes are denoted by a_l^+ and b_l^+ is considered first. When this wave encounters the clamped end, some reflected longitudinal, propagating transverse and evanescent transverse waves are created. Their magnitudes are denoted, respectively, a_l^-, b_l^-, $a_f^{p^-}$, $b_f^{p^-}$, $a_f^{e^-}$, and $b_f^{e^-}$. By applying the classical clamped boundary conditions (i.e., $\mu=\nu=\theta=0$) and recalling that magnitudes b are proportional to magnitudes a, three equations determine the magnitudes a_l^-, $a_f^{p^-}$, $a_f^{e^-}$ in terms of a_l^+;

$$\begin{pmatrix}a_l^-\\a_f^{p^-}\\a_f^{e^-}\end{pmatrix}=[r]_{sa}a_1^+. \tag{C1}$$

Now, applying (B2), relationships between the reflected energy flows and the incident energy flows carried by the incident longitudinal wave are determined as

$$\begin{pmatrix} P_l^- \\ P_f^- \end{pmatrix}_l = [R]_{22} \begin{pmatrix} P_l^+ \\ P_f^+ \end{pmatrix}_l . \tag{C2}$$

The subscript l following the vectors indicates that these energy flows are obtained for an incident longitudinal wave. Let us recall that in (C2) the unknowns are the coefficients of the matrix $[R]$. In a same way, another relationship is derived by considering an incident flexural propagating wave instead of a longitudinal wave. Then

$$\begin{pmatrix} P_l^- \\ P_f^- \end{pmatrix}_f = [R]_{22} \begin{pmatrix} P_l^+ \\ P_f^+ \end{pmatrix}_f . \tag{C3}$$

Both (C2) and (C3) provide all the material necessary to calculate the four coefficients of the matrix $[R]$. A numerical verification shows that for a undamped system (evanescent waves do not carry any energy flow), the sum of each column of $[R]$ is equal to one. This fact is relevant to a nondissipative end.

Finally, the reflection efficiencies for either longitudinal or flexural model are just the diagonal coefficients of the matrix $[R]$:

$$r_l = R_{1,1} \quad \text{and} \quad r_f = R_{2,2} . \tag{C4}$$

[1] R. H. Lyon, *Statistical Energy Analysis of Dynamical Systems: Theory and Applications* (MIT, Cambridge, 1975).

[2] E. H. Dowell and Y. Kubota, ''Asymptotic Modal Analysis and Statistical Energy Analysis of Dynamical Systems,'' J. Appl. Mech. **52**, 949–957 (1985).

[3] S. M. Doherty and E. H. Dowell, ''Experimental study of asymptotic modal analysis applied to a rectangular plate with concentrated masses,'' J. Sound Vib. **170**, 671–681 (1994).

[4] J. L. Guyader, ''Modal sampling method for the vibration study of systems of high modal density,'' J. Acoust. Soc. Am. **88**, 2269–2276 (1990).

[5] V. D. Belov, S. A. Rybak, and B. D. Tartakovski, ''Propagation of vibrational energy in absorbing structures,'' J. Sov. Phys. Acoust. **23**, 115–119 (1977).

[6] E. Luzzato, ''Approximations and solutions of the vibration energy density equations in beams,'' Inter Noise 1991.

[7] D. J. Nefske and S. H. Sung, ''Power flow finite element analysis of dynamic systems: Basic theory and applications to beams,'' Statistical Energy Analysis **3**, 47–54 (1987).

[8] J. C. Wohlever and R. J. Bernhard, ''Mechanical energy flow models of rods and beams,'' J. Sound Vib. **153**, 1–19 (1992).

[9] O. M. Bouthier and R. J. Bernhard, ''Simple models of energy flow in vibrating membranes,'' J. Sound Vib. **182**, 129–147 (1995).

[10] O. M. Bouthier and R. J. Bernhard, ''Simple models of the energetics of transversely vibrating plates,'' J. Sound Vib. **182**, 149–164 (1995).

[11] L. S. D. Morley, ''Elastic waves in a naturally curved rod,'' Q. J. Mech. Appl. Math. **14**, (1961).

[12] L. Cremer and M. Heckl, *Structure Born Sound: Structural Vibrations and Sound Radiation at Audio Frequencies* (Springer-Verlag, Berlin, 1973).

[13] Y. Lase and L. Jezequel, ''Analysis of a dynamic system based on a new energetic formulation,'' International Congress on Intensity Techniques, Senlis CETIM (1990).

[14] M. Djimadoum and L. Guyader, ''Prediction of coupled beam energy with the equation of diffusion-boundary, excitation and coupling conditions,'' International Congress on Intensity Techniques, Senlis CETIM (1993).

[15] M. N. Ichchou, ''Formulations énergétiques pour l'étude des moyennes et hautes fréquences des systèmes: Théorie et applications,'' Thèse École Centrale de Lyon No. 96-10.

[16] A. Le Bot and L. Jezequel, ''Energy methods applied to transverse vibrations of beams,'' International Congress on Intensity Techniques, Senlis CETIM (1993).

Elasto-acoustics of a two-dimensional thin strip by a hybrid method

Michael El-Raheb
The Dow Chemical Company, Midland, Michigan 48674

(Received 18 December 1996; accepted for publication 12 April 1997)

Elastic response of a two-dimensional thin strip coupled to an acoustic fluid is treated. Frequency response of the strip is expressed using transfer matrices of strip segments, and acoustic pressure of the fluid is expressed using a constant source density applied over each segment. Acoustic pressure loading the strip and continuity of elastic and acoustic velocities at midpoints of segments serve to couple the elastic and acoustic fields. Transmission loss (TL) of a baffled simply supported flat strip excited acoustically by diffuse sound is compared to TL from a modal solution, and TL measured across a rectangular plate having the same thickness and material properties. The unexpected close agreement between TL from the two-dimensional (2-D) strip model and measured TL is explained by relating strip and plate's analytical solutions of frequency response, and acoustic impedance from a prescribed oscillatory motion. © *1997 Acoustical Society of America.* [S0001-4966(97)02208-X]

PACS numbers: 43.40.Dx, 43.40.Rj [CBB]

INTRODUCTION

The problem of noise transmitted by the interaction of an elastic medium in contact with an acoustic fluid is finding an expanding number of applications because of the demand for a quieter environment in urban developments like transportation, housing, and mechanical appliances. A large body of work that treats this coupled structural acoustic problem exists in the literature.[1–16] Simplified analytical models of a flat two-dimensional (2-D) strip vibrating in an acoustic medium were among the earliest attempts to quantify transmission loss (TL), since the simplified equations yielded analytical solutions.[1,2] As the structures to be modeled became more complex, prediction of TL required the development of general purpose algorithms which systematically discretize the structure and acoustic fluid into finite elements. For external acoustic radiation from a vibrating structure, the large number of fluid elements needed for an accurate description of the acoustic pressure led to the development of the boundary element method which limits the number of acoustic elements to those over the boundary of the structure in contact with the fluid.[16,17] This method reduces the number of unknowns, but at the cost of destroying the banded nature of matrices, compromising computational efficiency.

General purpose computer programs[18–21] rely on several related computational schemes which differ principally in the simulation of the acoustic field. One commonly used scheme considers the jump in acoustic pressure Δp across the thin segment of plate or shell as the unknown. This yields a Green's function with a singularity higher in order than the integrable singularity in the boundary element method with pressure as the unknown. To address this difficulty, the singularity is reduced by requiring Δp to be orthogonal to an assumed smooth functional.[16] All programs adopt the modal expansion method to simulate frequency response of the thin elastic structure by discretizing it into finite elements. The displacement vector at nodal points is expressed as a superposition of eigenfunctions of the undamped structure in vacuum. These modes provide an *a priori* description of the spatial dependence needed to calculate modal acoustic pressure and its gradient. Substituting modal quantities into the equations of dynamic equilibrium and conditions of continuity of elastic and acoustic velocities at the structure–fluid interface, then exploiting orthogonality of the eigenfunctions determines simultaneous equations in the constant coefficients of the modal expansion and Δp. The steps of computing modal pressure and modal orthogonality require the evaluation of a double integral in a 2-D case like a strip and a quadruple integral in a three-dimensional (3-D) case like a plate.[15] Although efficient computational algorithms were devised to evaluate these integrals for a rectangular panel,[13] the calculation becomes demanding for more general geometries. Moreover, one limitation of the modal expansion method is high modal density of a composite structure when response is needed at frequencies producing short wavelengths. Another limitation is the inadequacy of in-vacuum eigenfunctions to problems where fluid density is comparable to structural areal density, or when material viscoelasticity produces substantial dissipation. These conditions strongly modify resonant frequencies and mode shapes.

This work develops a hybrid method treating the elastoacoustic interaction of a thin flat or curved strip surrounded by an acoustic fluid. The strip is divided into small segments with size a fraction of the shortest acoustic wavelength considered. The state vector **S** of forces and displacements at one end of a segment is related to **S** at the other end by a transfer matrix **T** according to an analytical solution of the homogeneous elastodynamic equations in the frequency domain. This method overcomes the difficulty from high modal density and viscoelastic dissipation since **T** is a frequency-dependent complex matrix accounting for complex valued material properties. The acoustic pressure and its gradient are analyzed using a 2-D Green's function and boundary elements. Pressure, assumed constant over a segment, is the sum of influence coefficients multiplied by an unknown source density vector. Concentrated forces from pressure integrated over a segment are assumed to act only at interfaces

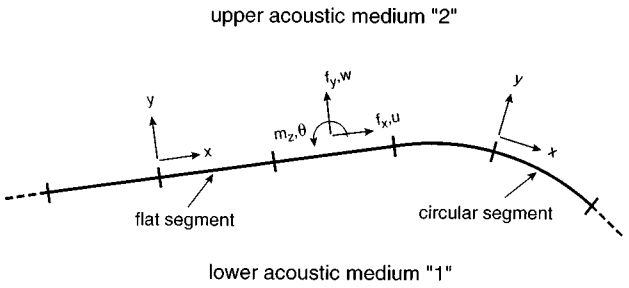


FIG. 1. The 2-D thin strip made of flat and circular segments.

of segments. Expressing continuity of elastic and acoustic velocities at midpoints of a segment completes the coupling between fluid and strip.

Section I develops the hybrid model of the 2-D strip. Section II applies the model to a baffled flat strip and compares TL with a modal solution and with measured data across a rectangular panel with the same thickness and material properties. Section III shows how a 2-D strip model differs from a 3-D plate model by analyzing the separate structural and acoustic responses. The difference in structural response is evaluated by analytical solutions for a simply supported strip and for a plate, each excited by a uniform pressure with periodic time dependence. The difference in acoustic response is evaluated by comparing acoustic pressure produced by a strip with that produced by a plate, prescribing the same oscillatory motion.

I. CONSTRUCTING THE HYBRID METHOD

The hybrid method is made up of three steps. The first step uses equations of the strip's structural response to create analytical transfer matrices of the strip's segments. Continuity at interfaces of segments yields a global matrix relating the ensemble of state vectors at all interfaces to the external excitation, from known point forces and net acoustic pressure, applied at these interfaces. The second step uses influence coefficients to relate source density applied over each segment to acoustic pressure and its normal gradient. The third step couples elastic and acoustic fields by requiring continuity of acoustic and elastic velocities at the center of each segment.

A. Analysis of structural response

The strip is divided into short flat or curved segments. Let $\mathbf{S}(x)$ be the state vector of forces $\mathbf{f}$ and displacements $\mathbf{g}$ at a point on the segment (see Fig. 1) referred to a local orthogonal coordinate system (x,y) with origin at one end of the segment where x is tangent to the segment:

$$\mathbf{S}(x)=\{\mathbf{f},\mathbf{g}\}=\{f_x,f_y,m_z;u,w,\theta\}^T, \tag{1}$$

where $0\leqslant x\leqslant l$ and l is the intrinsic length of a segment. Solving the homogeneous dynamic equations of equilibrium with periodic motions in time yields

$$\mathbf{S}(x)=\mathbf{B}(x)\mathbf{C}, \tag{2}$$

where $\mathbf{B}(x)$ is the solution matrix in terms of primitives of the differential operator and $\mathbf{C}$ is a vector of constant coefficients. Evaluating $\mathbf{S}$ at $x=0$ and $x=l$ then eliminating $\mathbf{C}$ produces the transfer matrix of a segment $\mathbf{T}$

$$\mathbf{S}(l)=\mathbf{B}(l)\mathbf{B}^{-1}(0)\mathbf{S}(0)\equiv\mathbf{T}(l\rightarrow 0)\mathbf{S}(0). \tag{3}$$

Derivations of $\mathbf{B}(x)$ and $\mathbf{T}$ for flat and circular segments are found in Appendix A. Continuity of $\mathbf{S}$ at interfaces of segments and constraints at the ends of the strip produces the global tridiagonal transfer matrix $\mathbf{T}_G$ relating the global state vector $\mathbf{S}_G$ to the external concentrated force vector $\mathbf{F}_{G0}$ at $(N+1)$ interfaces of segments

$$\mathbf{T}_G\mathbf{S}_G=\mathbf{F}_{G0},$$

$$\mathbf{S}_G=\{\mathbf{S}_1,\mathbf{S}_2,\ldots,\mathbf{S}_{N+1}\}^T\equiv\{\mathbf{f}_1,\mathbf{g}_1,\mathbf{f}_2,\mathbf{g}_2,\ldots,\mathbf{f}_{N+1},\mathbf{g}_{N+1}\}^T, \tag{4}$$

$$\mathbf{F}_{G0}=\{\mathbf{F}_{10},\mathbf{F}_{20},0,\mathbf{F}_{30},0,\ldots,0,\mathbf{F}_{N+1,0}\}^T,$$

$$\mathbf{T}_G=\begin{bmatrix}
-\mathbf{I} & \mathbf{Z}_1 & \mathbf{0} & \mathbf{0} & & & & & & \\
\mathbf{t}_{11}^1 & \mathbf{t}_{12}^1 & -\mathbf{I} & \mathbf{Z}_2 & & & & & & \\
\mathbf{t}_{21}^1 & \mathbf{t}_{22}^1 & \mathbf{0} & -\mathbf{I} & \mathbf{0} & \mathbf{0} & & & & \\
\mathbf{0} & \mathbf{0} & \mathbf{t}_{11}^2 & \mathbf{t}_{12}^2 & -\mathbf{I} & \mathbf{Z}_3 & & & & \\
 & & \mathbf{t}_{21}^2 & \mathbf{t}_{22}^2 & \mathbf{0} & -\mathbf{I} & & & & \\
 & & & & \vdots & \vdots & & & & \\
 & & & & & & \mathbf{t}_{11}^N & \mathbf{t}_{12}^N & -\mathbf{I} & \mathbf{Z}_N \\
 & & & & & & \mathbf{t}_{21}^N & \mathbf{t}_{22}^N & \mathbf{0} & \mathbf{I} \\
 & & & & & & \mathbf{0} & \mathbf{0} & -\mathbf{I} & \mathbf{Z}_{N+1}
\end{bmatrix}. \tag{4a}$$

In (4a)

$$\begin{bmatrix}\mathbf{t}_{11}^j & \mathbf{t}_{12}^j\\ \mathbf{t}_{21}^j & \mathbf{t}_{22}^j\end{bmatrix}\equiv\mathbf{T}_j$$

are the four submatrices forming the transfer matrix of the jth segment, $\mathbf{T}_j$, $\mathbf{Z}_j$ is a (3×3) impedance matrix connecting the jth interface to a fixed point in space, and $\mathbf{I}$ is the (3×3) unit matrix.

B. Analysis of acoustic response

A source density of constant spatial intensity σ_j is distributed along the jth strip segment. Acoustic pressure and its normal gradient at the central point of the ith segment with unit normal n_i are given by

$$p_i=\sum_{j=1}^{N} A_{ij}\sigma_j, \tag{5a}$$

$$\frac{\partial p_i}{\partial n_i}=\sum_{j=1}^{N} B_{ij}\sigma_j, \tag{5b}$$

where A_{ij} and B_{ij} are pressure and pressure gradient influence coefficients of the source density over the jth segment. In two dimensions,

$$A_{ij}=-\frac{1}{2}\int_0^{s_j}[Y_0(k_f r_{ij})-iJ_0(k_f r_{ij})]ds,\quad i=\sqrt{-1}, \tag{6a}$$

$$B_{ij}=\frac{k_f}{2}\int_0^{s_j}[Y_1(k_f r_{ij})-iJ_1(k_f r_{ij})]ds,$$

$$r_{ij}=((x_i-x_j)^2+(y_i-y_j)^2)^{1/2},\quad k_f=\frac{\omega}{c_f}, \tag{6b}$$

where c_f is speed of sound in the acoustic fluid, ω is radian frequency, s $(0\leqslant s\leqslant s_j)$ is an intrinsic coordinate along the jth segment, and J_n, Y_n are Bessel and Neuman functions. As $r\to 0$, $Y_0(k_f r)$ has a $\ln(r)$ integrable singularity while $Y_1(k_f r)$ has a $(1/r)$ singularity in the form of a Cauchy principal value. A_{ij} and B_{ij} are decomposed into

$$A_{ij}=\mathscr{R}_{ij}+\mathscr{S}_{ij},\quad B_{ij}=\mathscr{R}_{nij}+\mathscr{S}_{nij}, \tag{7}$$

where $\mathscr{R}$ and $\mathscr{S}$ are regular and singular parts of the integrals in (6). Expressions for A_{ij}, B_{ij}, $\mathscr{R}$, and $\mathscr{S}$ are derived in Appendix B for flat and curved segments.

C. Analysis of coupled response

Consider first the unbaffled case in which a single fluid without any acoustically isolated parts is in contact with both sides of a thin strip. The strip is divided into small segments. Let subscripts ''1'' and ''2'' refer to quantities on the ''lower'' and ''upper'' faces of a segment (see Fig. 1). Global dynamic equilibrium of all segments of the strip is expressed by (4)

$$\mathbf{T}_G\mathbf{S}_G=\mathbf{F}_{G_0}=\overline{\mathbf{p}}_2-\overline{\mathbf{p}}_1+\overline{\mathbf{p}}_{02}-\overline{\mathbf{p}}_{01}+\mathbf{F}_0. \tag{8}$$

Here, $\overline{\mathbf{p}}_2$ and $\overline{\mathbf{p}}_1$ are forces from acoustic pressure produced by motion of the strip and found by Eq. (5a) integrated over segments and acting at interfaces, $\overline{\mathbf{p}}_{02}$ and $\overline{\mathbf{p}}_{01}$ are known forces from acoustic pressure generated by external acoustic sources, and $\mathbf{F}_0$ are external point forces. Expressing $\overline{\mathbf{p}}$ in terms of influence coefficients

$$\overline{\mathbf{p}}_2-\overline{\mathbf{p}}_1=\overline{\mathbf{A}}_{e2,2}\sigma_2+\overline{\mathbf{A}}_{e2,1}\sigma_1-\overline{\mathbf{A}}_{e1,2}\sigma_2-\overline{\mathbf{A}}_{e1,1}\sigma_1, \tag{9}$$

where $\overline{\mathbf{A}}_{ek,l}$ is the extended pressure influence coefficient matrix along the kth face due to the source density vector σ_l along the lth face. Since the segment is thin,

$$\overline{\mathbf{A}}_{e2,2}=\overline{\mathbf{A}}_{e1,1}\quad\text{and}\quad\overline{\mathbf{A}}_{e2,1}=\overline{\mathbf{A}}_{e1,2} \tag{10}$$

which simplifies (9) to

$$\overline{\mathbf{p}}_2-\overline{\mathbf{p}}_1=(\overline{\mathbf{A}}_{e2,2}-\overline{\mathbf{A}}_{e2,1})(\sigma_2-\sigma_1). \tag{11}$$

In $\overline{\mathbf{A}}_{ek,l}(i',j)$ all rows vanish except when $i'=2+(i-1)6$ in which case $A_{ek,l}(i',j)=A(i,j)\Delta l_i$ where Δl_i is the length of the ith segment. Substituting (11) in (8) determines $\mathbf{S}_G$

$$\begin{aligned}&\mathbf{S}_G=\mathbf{T}_G^{-1}\{(\overline{\mathbf{A}}_{e2,2}-\overline{\mathbf{A}}_{e2,1})(\sigma_2-\sigma_1)\}+\mathbf{S}_0,\\&\mathbf{S}_0=\mathbf{T}_G^{-1}\{\overline{\mathbf{p}}_{02}-\overline{\mathbf{p}}_{01}+\mathbf{F}_0\},\end{aligned} \tag{12}$$

where $\mathbf{S}_0$ is a state vector from external acoustic and mechanical excitation. Define the matrix $\mathbf{M}$ as

$$\mathbf{M}=\mathbf{T}_G^{-1}(\overline{\mathbf{A}}_{e2,2}-\overline{\mathbf{A}}_{e2,1}) \tag{13}$$

and $\mathbf{W}$ as the reduced version of $\mathbf{M}$ including only rows $5+(i-1)6$ which correspond to components w_k in $\mathbf{S}_k$. Then

$$\mathbf{w}=\mathbf{W}(\sigma_2-\sigma_1)+\mathbf{w}_0 \tag{14}$$

and $\mathbf{w}_0$ is the reduced vector of $\mathbf{S}_0$ including the same rows as $\mathbf{W}$.

Continuity of elastic and acoustic velocities requires that

$$\rho_f\omega^2\mathbf{w}=\frac{\partial\mathbf{p}}{\partial\mathbf{n}}, \tag{15}$$

where ρ_f is the density of the acoustic fluid. Invoking (5b) in (15) yields

$$\rho_f\omega^2\mathbf{w}_2=\mathbf{B}_{2,2}\sigma_2+\mathbf{B}_{2,1}\sigma_1+\mathbf{V}_{02}, \tag{16a}$$

$$-\rho_f\omega^2\mathbf{w}_1=-\mathbf{B}_{1,2}\sigma_2+\mathbf{B}_{1,1}\sigma_1-\mathbf{V}_{01}, \tag{16b}$$

$\mathbf{V}_{0k}$ is the normal gradient of $\mathbf{p}_{0k}$. Expressing $\mathbf{B}_{k,k}$ as the sum of a diagonal and a nondiagonal matrix

$$\begin{aligned}&\mathbf{B}_{2,2}=\mathbf{I}+\hat{\mathbf{B}}_{2,2},\\&\mathbf{B}_{1,1}=\mathbf{I}-\hat{\mathbf{B}}_{2,2},\end{aligned} \tag{17}$$

where $\mathbf{I}$ accounts for the self term and $\hat{\mathbf{B}}_{2,2}$ is the nondiagonal part, and noting that for a thin strip $\mathbf{w}_1=\mathbf{w}_2=\mathbf{w}$ reduces (16) to

$$\rho_f\omega^2\mathbf{w}=(\mathbf{I}+\hat{\mathbf{B}}_{2,2})\sigma_2+\mathbf{B}_{2,1}\sigma_1+\mathbf{V}_{02}, \tag{18a}$$

$$-\rho_f\omega^2\mathbf{w}=-\mathbf{B}_{2,1}\sigma_2+(\mathbf{I}-\hat{\mathbf{B}}_{2,2})\sigma_1-\mathbf{V}_{01}. \tag{18b}$$

Expressing (18) in terms of $(\sigma_2-\sigma_1)$

$$\begin{aligned}\mathbf{w}=&\frac{1}{2\rho_f\omega^2}[\mathbf{I}-(\hat{\mathbf{B}}_{2,2}+\mathbf{B}_{2,1})(\hat{\mathbf{B}}_{2,2}-\mathbf{B}_{2,1})](\sigma_2-\sigma_1)\\&+\frac{1}{2\rho_f\omega^2}[-(\hat{\mathbf{B}}_{2,2}+\mathbf{B}_{2,1})(\mathbf{V}_{02}-\mathbf{V}_{01})+(\mathbf{V}_{02}+\mathbf{V}_{01})].\end{aligned} \tag{19}$$

Equating (14)–(19) determines a matrix equation in the unknown vector $(\sigma_2-\sigma_1)$. Specializing to a thin strip and expanding $\overline{\mathbf{A}}_{e2,2}$ as a Taylor's series in powers of a small distance along the normal reduces (13) and (19) to

$$\mathbf{M}=h\mathbf{T}_G^{-1}\hat{\mathbf{B}}_{e2,2}\quad\hat{\mathbf{B}}_{e2,2}=\frac{\partial}{\partial n}\overline{\mathbf{A}}_{e2,2}, \tag{20a}$$

$$\begin{aligned}\mathbf{w}=&\frac{1}{\rho_f\omega^2}\mathbf{B}_{2,2}(\sigma_2-\sigma_1)+\frac{1}{2\rho_f\omega^2}[-(\mathbf{I}+2\hat{\mathbf{B}}_{2,2})\\&\times(\mathbf{V}_{02}-\mathbf{V}_{01})+(\mathbf{V}_{02}+\mathbf{V}_{01})],\end{aligned} \tag{20b}$$

where h is the thickness of the strip.

Consider next the case of a baffled strip in which the fluid in contact with the upper face is acoustically isolated from the fluid in contact with the lower face. Then in place of (13) we have

$$\overline{\mathbf{A}}_{e2,1}=0\Rightarrow\mathbf{M}=\mathbf{T}_G^{-1}\overline{\mathbf{A}}_{e2,2}. \tag{21}$$

In (19) $\mathbf{B}_{2,1}\equiv\mathbf{B}_{1,2}=0$ simplifying it to

$$\begin{aligned}\mathbf{w}=&\frac{1}{2\rho_f\omega^2}[\mathbf{I}-\hat{\mathbf{B}}_{2,2}\hat{\mathbf{B}}_{2,2}](\sigma_2-\sigma_1)+\frac{1}{2\rho_f\omega^2}\\&\times[-\hat{\mathbf{B}}_{2,2}(\mathbf{V}_{02}-\mathbf{V}_{01})+(\mathbf{V}_{02}+\mathbf{V}_{01})].\end{aligned} \tag{22}$$

Furthermore, if the strip is flat then $\hat{\mathbf{B}}_{2,2}=0$ so (22) yields

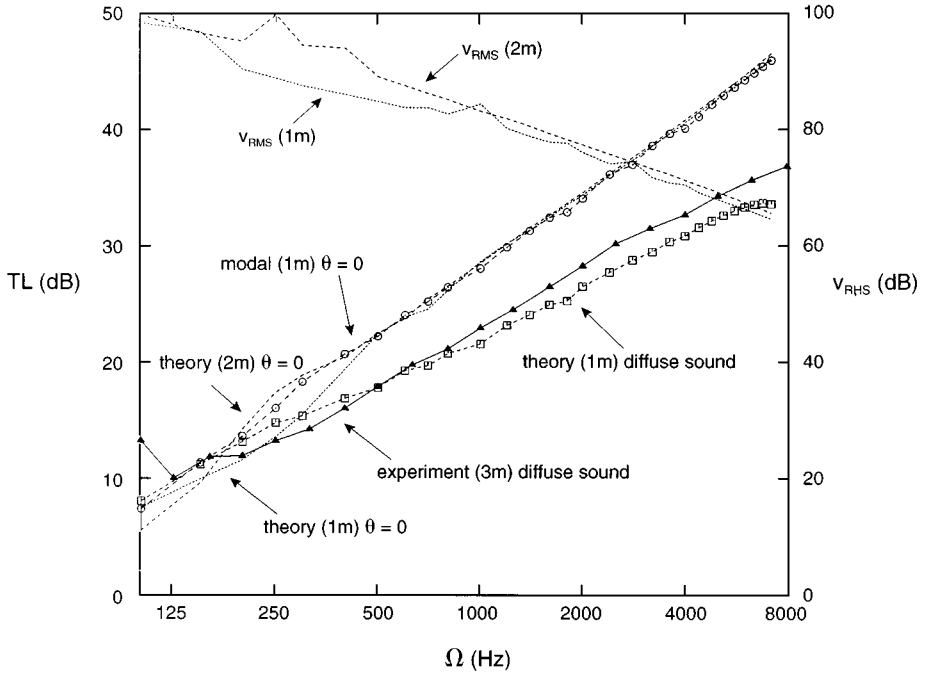


FIG. 2. Transmission loss (TL) of 1 and 2 m panels: $h=3$ mm polycarbonate sheet. ▲ experiment (3 m), diffuse sound; ⊙ modal (1 m), $\theta=0$. ⊡ theory (1 m), diffuse sound.

$$\mathbf{w}=\frac{1}{2\rho_f\omega^2}\left[(\sigma_2-\sigma_1)+\mathbf{V}_{02}+\mathbf{V}_{01}\right], \tag{23}$$

which when combined with (14) reduces to equations in $\mathbf{w}$:

$$\left(\mathbf{W}-\frac{1}{2\rho_f\omega^2}\,\mathbf{I}\right)\mathbf{w}=\frac{1}{2\rho_f\omega^2}\,\mathbf{W}(\mathbf{V}_{02}+\mathbf{V}_{01})-\frac{1}{2\rho_f\omega^2}\,\mathbf{w}_0, \tag{24a}$$

$$\sigma_2=\rho_f\omega^2 w-\mathbf{V}_{02}, \tag{24b}$$

$$\sigma_1=\rho_f\omega^2\mathbf{w}+\mathbf{V}_{01}. \tag{24c}$$

II. RESULTS

The hybrid analysis derived in Sec. I B is now applied to a flat baffled strip with simply supported edges. The strip is 2-m long and 3-mm thick, made of polycarbonate with the following material properties

$$E=2.3\times10^{10}\ \text{dyn/cm}^2,\quad \rho=1.2\ \text{g/cm}^3,$$

where (E,ρ) are Young's modulus and density. The strip is excited on the lower face by plane waves with incident pressure given by

$$p_{01}=P_0 e^{ik_f(x\sin\theta+y\cos\theta)}, \tag{25}$$

where P_0 is the amplitude of the plane wave and θ is the incident angle. At $\theta=0$, the plane wave is at normal incidence, while at $\theta=\pi/2$ it is at grazing incidence. Transmission loss (TL) is defined as transmitted power along the upper face divided by incident power along the lower face:

$$\text{TL}^{\theta}=\text{Re}\left\{\int_0^l (pv^*)dx\right\}\Big/\,(P_0^2 l_x\cos\theta/\rho_f c_f), \tag{26}$$

where v^* is the complex conjugate of acoustic velocity

$$v^*=\left(\frac{1}{i\rho_f\omega}\frac{\partial p}{\partial n}\right)^*.$$

TL from diffuse sound is defined as

$$\text{TL}^D=\int_0^{\pi/2}(\text{TL}^{\theta})\cos\theta\,d\theta. \tag{27}$$

Figure 2 shows TL^0 versus circular frequency Ω of the baffled plate at zero incidence and from diffuse sound. TL^0 agrees with that from the modal solution in Appendix C (see also Ref. 14). The TL^D agrees with data from an experiment performed on a rectangular panel with dimensions 4×3 m.[22] The fact that the experimental line is within 2 dB of the present analysis in spite of its reduced dimensionality can be understood by comparing the decoupled structural and acoustic response of strip and plate.

III. STRUCTURAL RESPONSE OF STRIP AND PLATE

Consider a thin simply supported flat strip excited by a uniform pressure with periodic time dependence and frequency ω. The dynamic equation is

$$\frac{Eh^3}{12(1-\nu^2)}\,w''''-\rho h\omega^2 w=P_0,\quad 0\leqslant x\leqslant l_x, \tag{28}$$

where (l_x,h) is the strip length and thickness, (E,ρ,ν) are the Young's modulus, mass density, and Poisson ratio, w is the transverse displacement, P_0 is the amplitude of applied pressure and $()'$ is the derivative with respect to axial coordinate x. The exact solution of (28) can be determined by the method of ''variation of parameters.'' [23] Equation (28) is recast in the form

$$w''''-k^4w=P_0/D,$$

$$D=\frac{Eh^3}{12(1-\nu^2)},\quad k^4=\frac{\omega^2}{c_0^2r_g^2},\tag{29}$$

$$c_0^2=\frac{E}{\rho(1-\nu^2)},\quad r_g^2=h^2/12.$$

The general solution of (29) assumes the form

$$w(x)=C_1(x)\sin kx+C_2(x)\cos kx+C_3(x)\sinh kx+C_4(x)\cosh kx.\tag{30}$$

Evaluating $w'(x)$, $w''(x)$, $w'''(x)$ from (30) and setting terms including $C_i'(x)$ to zero determines three equations in $C_i'(x)$. Substituting $w''''(x)$ and $w(x)$ in (29) determines the fourth equation in $C_i'(x)$. Solving for $C_i'(x)$ then integrating once yields the homogeneous and particular solutions in the form

$$w(x)=C_{01}\sin kx+C_{02}\cos kx+C_{03}\sinh kx+C_{04}\cosh kx+\frac{P_0}{2Dk^3}\int_0^x[-\sin k(x-\xi)+\sinh k(x-\xi)]d\xi.\tag{31}$$

For simple supports at both ends,

$$w(0)=w(l_x)=w''(0)=w''(l_x)=0.\tag{32}$$

Substituting (32) in (31) determines C_{0i} and $w(x)$ in the form

$$w(x)=\frac{P_0}{2Dk^4}[f_p(x)+f_e(x)],$$

$$f_p(x)=\frac{\sin kx-\sin kl_x+\sin k(x-l_x)}{\sin kl_x},\tag{33}$$

$$f_e(x)=\frac{\sinh kx-\sinh kl_x+\sinh k(x-l_x)}{\sinh kl_x}.$$

Since $Dk^4=\rho h\omega^2$ it follows that mobility (w/P_0) is inversely proportional to ω^2. Defining rms velocity as

$$v_{\rm rms}=20\log_{10}\left[\frac{1}{l}\int_0^l\dot{w}^2(x)dx\right]^{1/2}\tag{34}$$

then substituting (33) in (34) yields

$$v_{\rm rms}=-20\log_{10}(\rho h\omega)+\text{const.}$$

Now, to contrast, consider a simply supported rectangular plate with thickness, material properties, and excitation the same as the strip's. The dynamic equation is

$$D\nabla^4w-\rho h\omega^2w=P_0,\quad 0\leqslant x\leqslant l_x,\quad 0\leqslant y\leqslant l_y,$$

$$\nabla^4\equiv(\partial_{xx}+\partial_{yy})^2\tag{35}$$

with boundary conditions

$$w(0,y)=w(l_x,y)=0,$$

$$(w_{xx}+\nu w_{yy})=0\quad\text{at } x=0,l_x\ \ \forall y,$$

$$w(x,0)=w(x,l_y)=0\tag{36}$$

$$(w_{yy}+\nu w_{xx})=0\quad\text{at } y=0,l_y\ \ \forall x.$$

Boundary conditions along $y=0,l_y$ suggest a solution in the form

$$w(x,y)=\sum_{n=1}^{\infty}w_n(x)\sin k_ny,\quad k_n=n\pi/l_y,\tag{37}$$

where n is the integer wave number along y. Substituting (37) in (35) while enforcing orthogonality of $\sin(k_yy)$ yields equations in $w_n(x)$

$$w_n''''-2k^2w_n''+k_n^4w_n-k^4w_n=\frac{4}{l_yk_n}P_0,\tag{38}$$

where k^4 is defined in (29) and $(\,)'$ is derivative $w\cdot r\cdot t\cdot x$. Equation (38) admits a solution in the form

$$w_n(x)\propto e^{\hat{k}_xx},\tag{39}$$

which when substituted in (38) yields the dispersion relation

$$\hat{k}_x^4-2k_n\hat{k}_x^2+(k_n^4-k^4)=0\tag{40a}$$

with solutions

$$-\hat{k}_{x_1}^2=k_{x_1}^2=(m\pi/l_x)^2=k^2-k_n^2$$

$$\Rightarrow\text{primitives: } \sin k_{x1}x,\cos k_{x1}x,\tag{40b}$$

$$\hat{k}_{x_2}^2=k_{x_2}^2=k^2+k_n^2\Rightarrow\text{primitives: } \sinh k_{x2}x,\cosh k_{x2}x.$$

For $k_{x_1}^2>0$, k_n must be in the range $\pi/l_y\leqslant k_n\leqslant k$. The general solution to $w_n(x)$ is determined adopting the same steps that led to the solution of the strip given by Eq. (33):

$$w_n(x)=\frac{2P_0}{Dk_nk^2l_y}\left\{\frac{\sin k_{x1}x-\sin k_{x1}l_x+\sin k_{x1}(l_x-x)}{k_{x_1}^2\sin k_{x1}l_x}+\frac{\sinh k_{x2}x-\sinh k_{x2}l_x+\sinh k_{x2}(l_x-x)}{k_{x_2}^2\sinh k_{x2}l_x}\right\}.\tag{41}$$

The first and second terms in (41) produce propagating and evanescent waves, respectively. Remote from resonance, terms in (37) with largest contribution to response result from the the first term in (41) when its denominator is smallest. The part of the denominator which varies with k_n is

$$d_n\equiv k_nk_{x_1}^2=k_n(k^2-k_n^2),$$

d_n is largest when

$$\frac{d(d_n)}{dk_n}=k^2-3k_n^2=0\Rightarrow k_n=k/\sqrt{3}\Rightarrow n=\text{integer}\left(\frac{kl_y}{\pi\sqrt{3}}\right)$$

and d_n is smallest at the two bounds on k_n such that $k_{x1}>0$:

lower bound: $n_{\min}=1\Rightarrow k_{n1}=\pi/l_y$,

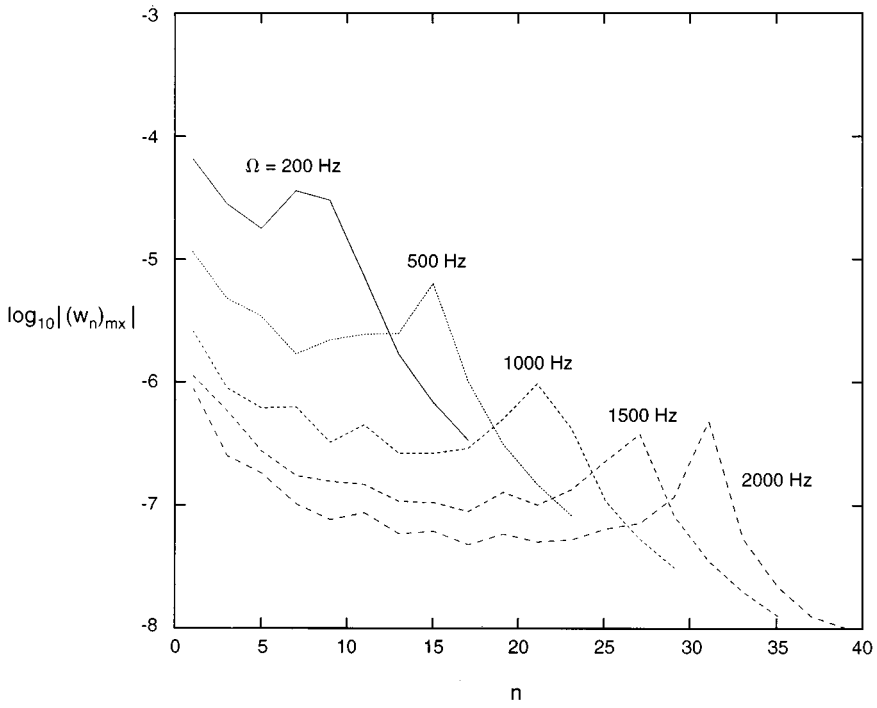


FIG. 3. Maximum modal amplitude versus wave number n along y of the square simply supported plate $l_x=l_y=1$ m, $h=0.3$ cm.

$$m_{\max}=\frac{l_x}{\pi}\sqrt{k^2-k_{n1}^2},$$

upper bound: $n_{\max}=\text{integer}\ (kl_y/\pi)\Rightarrow k_{n2}=n_{\max}\pi/l_y$,

$$m_{\min}=\frac{l_x}{\pi}\sqrt{k^2-k_{n2}^2}. \tag{42}$$

Consequently, the two dominant modes in the plate's response are $(m_{\max},1)$ and $(m_{\min},n_{\max})$ where bounds on these wave numbers are defined above. Substituting k_{n1} into (41) assuming $k^2\gg(\pi/l_y)^2$ yields

$$w_{n1}(x)\simeq\frac{2P_0}{\pi Dk^4}\,[f_p(x)+f_e(x)], \tag{43}$$

where $f_p(x)$ and $f_e(x)$ are defined in (33). Substituting k_{n2} in (41) and taking the limit as $k_n\to k$ leads to

$$w_{n2}=\frac{P_0kl_x^2}{Dk^4l_y}\left\{\frac{x(l_x-x)}{l_x^2}+\frac{\sinh\sqrt{2}kx-\sinh\sqrt{2}kl_x+\sinh\sqrt{2}k(l_x-x)}{k^2l_x^2\sinh\sqrt{2}kl_x}\right\}. \tag{44}$$

The ratio between w_{n1} in (43) and w in (33) is $(4/\pi)$, a difference of 2 dB, consistent with the difference observed in Fig. 2 between experimental and theoretical TL^D lines. Figure 3 plots $\log_{10}|\max(w_n(x))|$ against n with Ω as parameter. Note peak response at $n_{\min}$ and $n_{\max}$ and the sharp dip when $n>n_{\max}$ when waves become evanescent.

Figure 4 plots v_{rms} of strip and square plate for $\theta=0$ and $\theta=\pi/2$. When $\theta=0$, v_{rms} of strip and plate almost coincide. When $\theta=\pi/2$, v_{rms} of strip and plate have the same bounds but corresponding resonant peaks are shifted. This comparison demonstrates that while frequency response of a plate differs from the strip's by as much as 25% (2 dB), for understanding the overall elastic behavior, the strip forms a very close approximation to the more difficult plate.

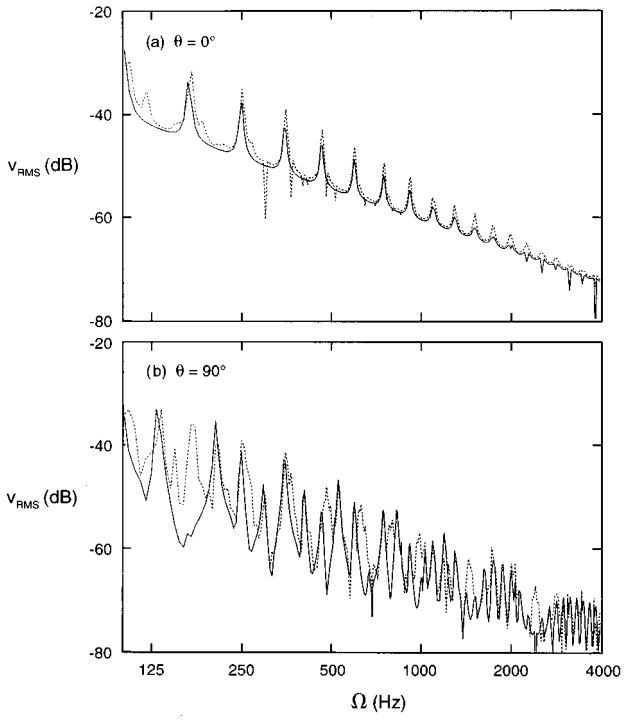


FIG. 4. The v_{rms} response of strip (———) and plate (-------) excited by a plane wave at incidence (a) $\theta=0°$, (b) $\theta=90°$.

IV. ACOUSTIC RESPONSE OF STRIP AND PLATE

Consider first a flat baffled strip with prescribed velocity $v(x)$ along its length l_x where x is the axial coordinate with origin at one end of the strip

$$v(x)=v_0\ \sin(m\pi x/l_x). \tag{45}$$

Acoustic response is measured by acoustic impedances Z at a point (x_0,z_0) on and remote from the strip where z_0 is normal distance of a field point from the strip

$$Z(x_0,z_0)=p(x_0,z_0)/v_0. \tag{46}$$

From Appendix C Eq. (C4):

$$p(x_0,z_0)=\int_0^{l_x} G_2(x_0-\xi,z_0)\ \frac{\partial p_0(\xi)}{\partial n}\ d\xi$$

$$=-i\rho_f\omega v_0\int_0^{l_x} G_2(x_0-\xi,z_0)\sin(m\pi\xi/l_x)d\xi \tag{47a}$$

$$\Rightarrow Z(x_0,z_0)=-i\rho_f c_f k_f\int_0^{l_x} G_2(x_0-\xi,z_0)$$

$$\times\sin(m\pi\xi/l_x)d\xi. \tag{47b}$$

Now by contrast consider a flat baffled square plate in the (x,y) plane with prescribed velocity $v(x,y)$ over its surface where the origin is at one corner of the plate

$$v(x,y)=v_0\ \sin(m\pi x/l_x)\sin(n\pi y/l_y) \tag{48}$$

with acoustic impedance at a point (x_0,y_0,z_0) given by

$$Z(x_0,y_0,z_0)=p(x_0,y_0,z_0)/v_0. \tag{49}$$

Acoustic pressure is expressed as

$$p(x_0,y_0,z_0)=\int_0^{l_y}\int_0^{l_x} G_3(x_0-\xi,y_0-\eta,z_0)$$

$$\times\frac{\partial p_0(\xi,\eta)}{\partial n}\ d\xi\ d\eta, \tag{50a}$$

where $G_3(x_0-\xi,y_0-\eta,z_0)$ is the 3-D acoustic Green's function

$$G_3(x_0-\xi,y_0-\eta,z_0)=\frac{e^{ik_f r}}{r},$$

$$r=[(x_0-\xi)^2+(y_0-\eta)^2+z_0^2]^{1/2}$$

$$\Rightarrow Z(x_0,y_0,z_0)$$

$$=-i\rho_f c_f k_f\int_0^{l_y}\int_0^{l_x} G_3(x_0-\xi,y_0-\eta,z_0)$$

$$\times\sin(m\pi\xi/l_x)d\xi\ \sin(n\pi\eta/l_y)d\eta. \tag{50b}$$

The double integral in (50b) is decomposed into a regular and a singular part $\mathscr{R}$ and $\mathscr{S}$

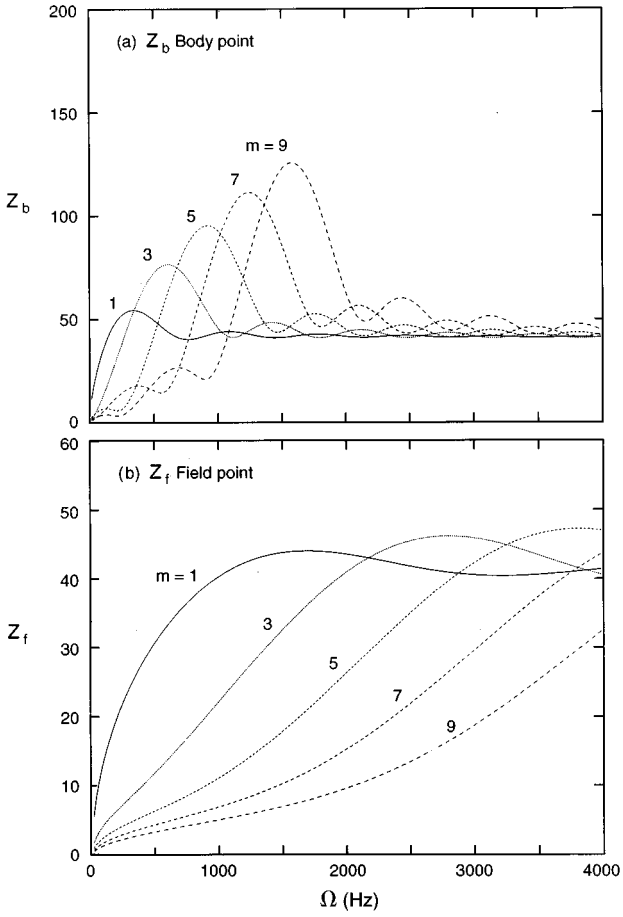


FIG. 5. Center point impedance of flat strip, (2-D) analysis. $l=1$ m, $x_b=0.5$ m, $x_f=0.5$ m, $y_f=1$ m. (a) Z_b Body point; (b) Z_f Field point.

$$\tilde{I}=\int_0^{l_y}\int_0^{l_x}\frac{e^{ik_f r}}{r}\sin\alpha\xi\ d\xi\ \sin\beta\eta\ d\eta=\mathscr{R}+\mathscr{S},$$

$$\alpha=\frac{m\pi}{l_x},\quad \beta=\frac{n\pi}{l_y}, \tag{51}$$

$$\mathscr{R}=\int_0^{l_y}\int_0^{l_x}\left(\frac{e^{ik_f r}}{r}-\frac{1}{r}\right)\sin\alpha\xi\ d\xi\ \sin\beta\eta\ d\eta, \tag{52a}$$

$$\mathscr{S}=\int_0^{l_y}\int_0^{l_x}\frac{1}{r}\sin\alpha\xi\ d\xi\ \sin\beta\eta\ d\eta$$

$$=\int_0^{l_y}\sin\beta\eta[\sin\alpha\xi\ \ln|\xi+r|]_{x_1}^{x_2}\ d\eta$$

$$-\alpha\int_0^{l_y}\sin\beta\eta[\cos\alpha\xi(\xi\ \ln|\xi+r|-r)]_{x_1}^{x_2}\ d\eta$$

$$+\alpha^2\int_0^{l_y}\int_0^{l_x}\sin\beta\eta\ \sin\alpha\xi(\xi\ \ln|\xi+r|-r)d\xi\ d\eta. \tag{52b}$$

Acoustic impedances are computed first for a 1-m-long strip at $x_0=0.5m$ on the strip and at a field point $(x_0,z_0)=(0.5,1)m$. The velocity prescribed along x is given by (45) for $m=1,3,5,7,9$. Figure 5(a) and (b) plot Z_b and Z_f at body and field points against frequency Ω. For each m, Z_b

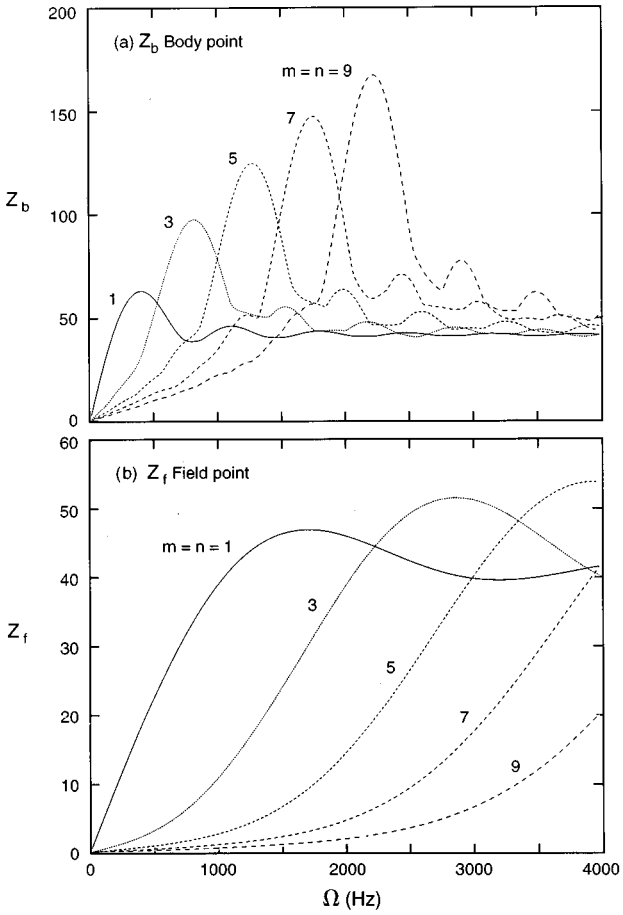


FIG. 6. Center point impedance of flat square plate, (3-D) analysis $l_x=l_z=1$ m, $(x_b,z_b)=(0.5\text{ m}, -0.5\text{ m})$, $(x_f,y_f,z_f)=(0.5\text{ m},1\ m, -0.5\text{ m})$. (a) Z_b Body point; (b) Z_f Field point.

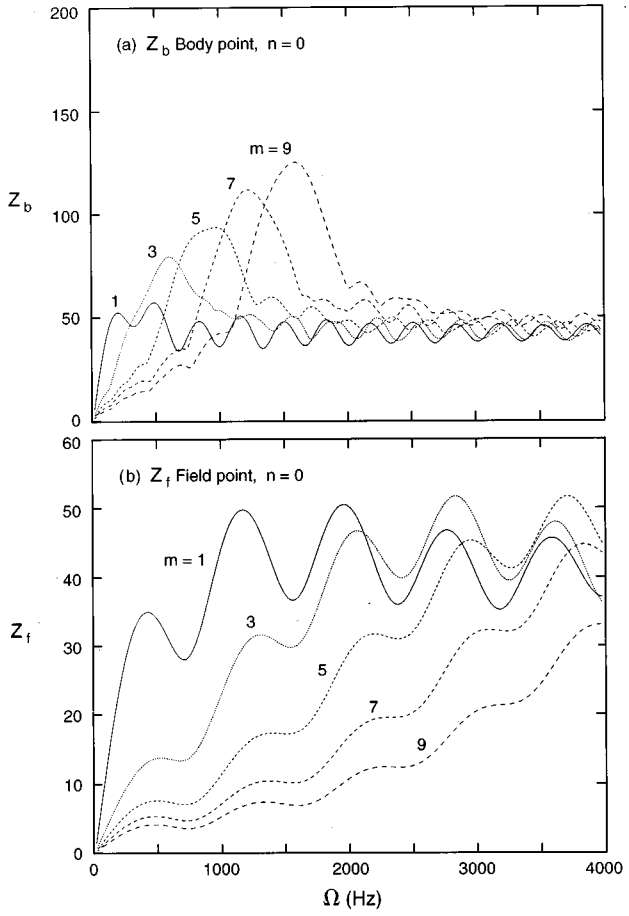


FIG. 7. Center point impedance of flat square plate with $n=0$, (3-D) analysis $l_x=l_z=1$ m, $(x_b,z_b)=(0.5\text{ m}, -0.5\text{ m})$, $(x_f,y_f,z_f)=(0.5\text{ m}, 1\text{ m}, -0.5\text{ m})$. (a) Z_b Body point; (b) Z_f Field point.

rises smoothly with Ω then peaks at coincidence

$$\Omega_{cm}=c_f\, m/(2l_x) \tag{53}$$

when wavelengths of sound and velocity distribution of the strip coincide. Above coincidence, Z_b falls then approaches the acoustic impedance of a plane wave $Z_0=\rho_f c_f$. As m increases, the slope of Z_b prior to coincidence falls while peak at coincidence rises. Above coincidence, the asymptote remains Z_0. The Z_f curve resembles that of Z_b except that peaks for all m are closer to the asymptote Z_0.

Acoustic impedances are computed next for a square plate with $l_x=l_y=1$ m, at $(x_0,y_0,z_0)=(0.5,0.5,0)$ m and $(x_0,y_0,z_0)=(0.5,0.5,1)$ m. Prescribed velocity is given by (48) with $m=n=1,3,5,7,9$. Figure 6(a) and (b) plot Z_b and Z_f versus frequency. Fixing (m,n), Z_b of the plate is similar in shape to that of the strip, while the peak at coincidence is higher. In general, there are two coincident frequencies

$$\Omega^{(1)}_{c(m,n)}=c_f\sqrt{m^2+n^2}/(2l_x),$$
$$\Omega^{(2)}_{c(m,n)}=c_f\sqrt{m^2+n^2}/(2l_y). \tag{54}$$

If $m=n$, and $l_x=l_y$ then the two coincidences coalesce yielding

$$\Omega_{cm}=\sqrt{2}c_f m/(2l_x).$$

Figure 7(a) and (b) plot Z_b and Z_f against Ω for $m=1,3,5,7,9$ independent of y, i.e., $v(x,y)=v_0\sin(m\pi x/l_x)$. Fixing m, Z_b and Z_f fluctuate about average lines corresponding to the strip's Z_b and Z_f shown in Fig. 5(a) and (b). This fluctuation has a period T_0 where

$$T_0\simeq c_f/\Delta r,$$
$$\Delta r=\sqrt{z_0^2+(l_x/2)^2}-z_0 \tag{55}$$

and Δr is the difference between the longest and the shortest distance from field point to plate. Appendix D derives T_0 for a line source extended in the x direction, undergoing periodic oscillations along z, i.e., $m=0$, and concludes that the fluctuation vanishes when $m>0$.

It was shown in Sec. III that the dominant modes in the frequency response of a plate are $(m,n)=(m_{\max},1)$ and $(m_{\min},n_{\max})$ where $m_{\max}$ and $n_{\max}$ are the highest wave numbers yielding propagating waves in the plate. Since these wave-number combinations also dominate the velocity distribution in (48), producing the acoustic field (Z_b,Z_f), then from the closeness of (Z_b,Z_f) for the plate and strip at $(m,n)=(m_{\max},1)$ when $m_{\max}>1$, the conclusion at the end of

this section may be broadened to state that for predicting acoustic response, too, the strip forms a very close approximation to the plate.

V. CONCLUSION

A hybrid method to compute the elastoacoustic interaction of a 2-D thin strip made of flat and curved segments is formulated. Analysis proceeds through elastic response using transfer matrices of segments and through acoustic response using 2-D boundary elements. The method succeeds even when modal density is high. Noteworthy results from the simulation of a baffled thin flat strip are

(1) TL from the hybrid method agrees both with that from a modal solution and also with that from experimental data when the plate is excited by diffuse sound

(2) Dominant modes in the plate's frequency response include $(m_{\max},1)$ and $(m_{\min},n_{\max})$ where (m,n) are wave numbers along (x,y).

(3) The dominant modes in plate response are identical to those of the strip.

(4) Acoustic response of the plate from its dominant elastic modes is comparable to acoustic response of the strip vibrating at the same highest elastic wave number.

(5) For the purpose of predicting acoustic response and TL, the strip is a very close approximation to the plate.

APPENDIX A: TRANSFER MATRIX OF A SEGMENT

Figure A1 shows a circular segment with end points 1 and 2. Let R_m be its radius of curvature, h the thickness, and $\mathbf{S}=\{f_x,f_y,m_z,u_x,u_y,\theta_z\}^T$ the state vector in local coordinates at station s along the arclength. Here, $\mathbf{S}$ satisfies the equilibrium equation

$$\frac{\partial f_x}{\partial s}-\frac{f_y}{R_m}=\rho h\ddot{u}_x, \tag{A1a}$$

$$\frac{\partial f_y}{\partial s}+\frac{f_x}{R_m}=\rho h\ddot{u}_y, \tag{A1b}$$

$$\frac{\partial m_z}{\partial s}+f_y=\rho\,\frac{h^3}{12}\,\ddot{\theta}_z, \tag{A1c}$$

where ρ is the density and $(\dot{})$ is the derivative with respect to time. The constitutive relations are

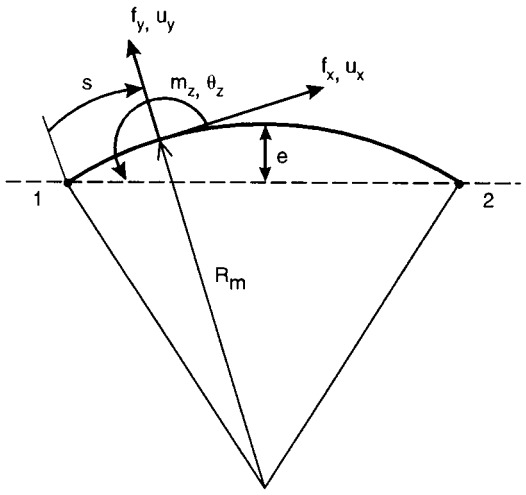


FIG. A1. Circular segment.

$$f_x=Eh\int_h \frac{\epsilon_x+y\kappa_z}{1+y/R_m}\,dy,$$

$$\epsilon_x=\frac{\partial u_x}{\partial s}-\frac{u_y}{R_m},\quad \kappa_z=\frac{\partial\theta_z}{\partial s}$$

$$\Rightarrow f_x=EZ_1\left(\frac{\partial u_x}{\partial s}-\frac{u_y}{R_m}\right)+EZ_2\,\frac{\partial\theta_z}{\partial s}, \tag{A2}$$

$$Z^*=\frac{R}{h}\ln\left(\frac{R_m/h+1/2}{R_m/h-1/2}\right)-1,$$

$$Z_1=h(1+Z^*),\quad Z_2=-R_mhZ^*,$$

$$f_y=\kappa Gh\left(\frac{\partial u_y}{\partial s}+\frac{u_x}{R_m}-\theta_z\right), \tag{A3}$$

$$m_z=EZ_2\left(\frac{\partial u_x}{\partial s}-\frac{u_y}{R_m}-R_m\,\frac{\partial\theta_z}{\sigma s}\right), \tag{A4}$$

where κ is shear constant. For periodic motions in time with frequency ω, (A1)→(A4) cast in matrix form are

$$\overline{\mathbf{B}}\,\frac{\partial\mathbf{S}}{\partial s}=\mathbf{BS},$$

$$\mathbf{B}=\begin{bmatrix} 0 & 1/R_m & 0 & -\rho h\omega^2 & 0 & 0\\ -1/R_m & 0 & 0 & 0 & -\rho h\omega^2 & 0\\ 0 & -1 & 0 & 0 & 0 & -\rho h^3\omega^2/12\\ 1/(EZ_1) & 0 & 0 & 0 & 1/R_m & 0\\ 0 & 1/(\kappa Gh) & 0 & -1/R_m & 0 & 1\\ 0 & 0 & -1/(EZ_2R_m) & 0 & -1/R_m^2 & 0 \end{bmatrix},$$

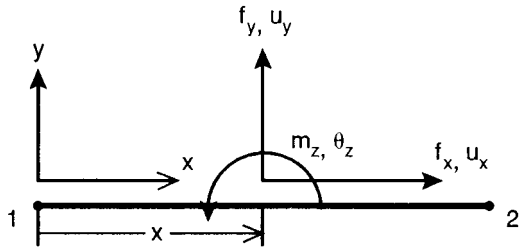


FIG. A2. Flat segment.

$$\overline{\mathbf{B}}=\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & Z_2/Z_1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1/R_m & 0 & 1 \end{bmatrix}. \tag{A5}$$

Equation (A5) admits a solution in terms of exponentials

$$\mathbf{S}(s)=\mathbf{B}_0\mathbf{e}^{\lambda s}. \tag{A6}$$

Substituting (A6) in (A5) yields

$$[\overline{\mathbf{B}}^{-1}\mathbf{B}-\lambda\mathbf{I}]\mathbf{B}_0=0. \tag{A7}$$

A nontrivial solution of (A7) yields the dispersion relation

$$\det|\overline{\mathbf{B}}^{-1}\mathbf{B}-\lambda\mathbf{I}|=0. \tag{A8}$$

In terms of the characteristic roots λ_i, the solution of (A6) takes the form

$$\mathbf{S}(s)=\mathbf{B}_0\begin{bmatrix} \backslash \\ \mathbf{e}^{\lambda s} \\ \backslash \end{bmatrix}\mathbf{C}, \tag{A9}$$

where $\mathbf{B}_0$ is the matrix of eigenvectors of (A8) and $\mathbf{C}$ is a vector of unknown constant coefficients. Evaluating (A9) at $s=0$ and $s=l$ then eliminating $\mathbf{C}$ yields the transfer matrix of the segment

$$\begin{aligned} &\mathbf{S}(l)=\mathbf{T}(l\rightarrow 0)\mathbf{S}(0), \\ &\mathbf{T}(l\rightarrow 0)=\mathbf{B}_0\begin{bmatrix} \backslash \\ \mathbf{e}^{\lambda l} \\ \backslash \end{bmatrix}\mathbf{B}_0^{-1}. \end{aligned} \tag{A10}$$

The derivation above can be applied to the flat segment (see Fig. A2) in the limit when $R_m\rightarrow\infty$. In this case, Z_1 and Z_2 in (A2) become

$$\lim_{R\rightarrow\infty} Z_1=h, \quad \lim_{R\rightarrow\infty} Z_2=0, \quad \lim_{R\rightarrow\infty} Z_2R_m=-h^2/12.$$

APPENDIX B: ACOUSTIC INFLUENCE COEFFICIENTS

Subscripts ''ij'' denoting field and source of coefficients A_{ij} and B_{ij} in (6a) and (6b) of the text will be omitted for brevity unless distinction of these element quantities is not obvious. The geometric parameters of flat and circular elements are shown in Fig. B1. As noted in Eq. (7), influence coefficients in (6a) and (6b) can be decomposed into a regular and a singular part. The singular behavior follows the logarithmic singularity in the Y_0 expansion as the field point approaches the source point. The pressure influence coefficient A in (6a) decomposes into

$$A=\mathscr{R}+\mathscr{S}, \tag{B1a}$$

where $\mathscr{R}$ and $\mathscr{S}$ are the regular and singular parts of the integral

$$\begin{aligned} \mathscr{R}&=-\frac{1}{2}\int_s\left[Y_0(k_fr)-iJ_0(k_fr)-\frac{2}{\pi}\ln(k_fr)/2\right]ds, \\ \mathscr{S}&=-\frac{1}{\pi}\int_s \ln(k_fr/2)ds \\ &=-\frac{1}{2\pi\cos\theta}\left\{2(\ln(k_f/2)-1)x+(x+b/2a)\ln\xi\right. \\ &\quad\left.+(4ac-b^2)^{1/2}\Big/ a\tan^{-1}\left[\frac{2ax+b}{(4ac-b^2)^{1/2}}\right]\right\}\Bigg|_{x_1}^{x_2}, \\ \xi&=ax^2+bx+c, \quad a=\sec^2\theta, \\ b&=-2[\overline{x}+\overline{y}\tan^2\theta(x_1\tan\theta-y_1)], \\ c&=\overline{x^2}+\overline{y^2}+(x_1\tan\theta-y_1)^2+2\overline{y}(x_1\tan\theta-y_1). \end{aligned} \tag{B1b}$$

θ is the angle of the jth flat element lying between points (x_1,y_1) and (x_2,y_2). The limits of the integral are from 0 to Δs, the length of the source element. The velocity influence coefficient decomposes into

$$B=\mathscr{R}_n+\mathscr{S}_n, \tag{B2a}$$

$$\begin{aligned} \mathscr{R}_n&=-\frac{k_f}{2}\int_s\left[Y_1(k_fr)-iJ_1(k_fr)+\frac{2}{\pi k_fr}\right]\sin\varphi_i\,ds, \\ \mathscr{S}_n&=\left\{f_3\ln\xi+f_4\tan^{-1}\left[\frac{2ax+b}{(4ac-b^2)^{1/2}}\right]\right\}\Bigg|_{x_1}^{x_2}, \\ f_3&=\frac{1}{2a\pi}f_1, \\ f_4&=\frac{2}{\pi(4ac-b^2)^{1/2}}[f_2-f_1(x_1+b/2a)], \\ f_1&=\frac{\sin(\theta_i-\theta_j)}{\cos^2\theta_j}, \quad f_2=r_{1j}\sin\varphi_{1i}/\cos\theta_j, \end{aligned} \tag{B2b}$$

where φ_i is the angle between r and the ith influenced element and φ_{1i} is the angle between r_{1j} and the ith influenced element (see Fig. B1).

Let (x_c,y_c) be the coordinates of the center of curvature c of the jth circular element with radius R. The element is confined within the arc defined by the angles χ_1 and χ_2 measured from an axis parallel to x and passing through (x_c,y_c) (see Fig. B1). The potential influence coefficient of a constant source distribution on element j at a field point i is

$$\begin{aligned} A&=-\frac{1}{2}\int_\chi[Y_0(k_fr)-iJ_0(k_fr)]R\,d\chi, \\ r&=[a-b\cos(\chi-\chi_i)]^{1/2}, \quad a=R^2+R_c^2, \quad b=2RR_c. \end{aligned} \tag{B3}$$

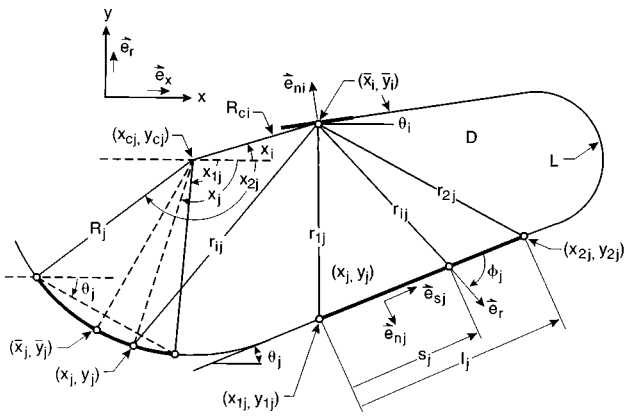


FIG. B1. Two-dimensional flat and circular boundary elements.

χ is the angular position of a general point (x,y) on the element and R_c is the radial distance between c and the central point of the ith influenced element with coordinates (x,y) while χ_i is its angular position (see Fig. B1). The limits of the integral are from χ_1 to χ_2. The integral (B3) can be decomposed into regular and singular parts in a manner similar to that adopted in the straight element.

$$A=\mathscr{R}+\mathscr{S}, \tag{B4a}$$

$$\mathscr{R}=-\frac{1}{2}\int_{\chi}\left[Y_0(k_f r)-iJ_0(k_f r)-\frac{2}{\pi}\ln(k_f r)/2\right]R\,d\chi,$$

$$\mathscr{S}=-\frac{R}{\pi}\left\{\chi\,\ln(k_f r)+\left[\psi\,\ln\,\psi-\psi-\sum_n\frac{2^{2n-1}B_n}{n(2n+1)!}\,\psi^{2n+1}\right]\right\}_{\chi}, \tag{B4b}$$

where B_n are the Bernoulli numbers and $\psi=(\chi-\chi_i)$. In the case where $a=b$, $\mathscr{S}$ takes the form

$$\mathscr{S}=-\frac{R}{\pi}\left\{\psi\,\ln(\sqrt{0.5}k_f R)+\frac{\psi}{2}\ln(1-\cos\,\psi)-2\int_{\eta_1}^{\eta_2}\frac{\sin^{-1}\,\eta}{\eta}\,d\eta\right\}_{\chi}, \tag{B4c}$$

where $\eta=\sin^{-1}(\chi/2)$, $a=2R^2$ and the bracket is evaluated at the limits $(\chi_1-\chi_i)$ and $(\chi_2-\chi_i)$. The velocity influence coefficient X in the x direction can be decomposed into regular and singular parts:

$$X=\mathscr{R}_x+\mathscr{S}_x,$$

$$\mathscr{R}_x=-\frac{k_f}{2}\int_{\chi}\left[Y_1(k_f r)-iJ_1(k_f r)+\frac{2}{\pi k_f r}\right]\frac{(x-\overline{x})}{r}\,R\,d\chi, \tag{B5a}$$

$$\mathscr{S}_x=\frac{1}{\pi}\left\{\frac{C_1}{b}\ln\left|a-b\,\cos\,\chi\right|-\frac{C_2}{b}\,\psi+\frac{2}{(a^2-b^2)^{1/2}}\times(C_0+aC_2/b)\tan^{-1}\frac{(a^2-b^2)^{1/2}\tan(\psi/2)}{a-b}\right\}_{\chi},$$

$$C_0=x_c-\overline{x},\quad C_1=-R\,\sin\,\chi_i,\quad C_2=R\,\cos\,\chi_i, \tag{B5b}$$

where a and b are as previously defined [following Eq. (B3)]. $\mathscr{S}_x$ in Eq. (B5b) is evaluated at the limits $(\chi_1-\chi_i)$ and $(\chi_2-\chi_i)$. A similar decomposition applies to the velocity influence coefficient Y in the y direction with $(y-\overline{y})$ replacing $(x-\overline{x})$ in the regular part of Eqs. (B5) and C_0, C_1, C_2 replacing C_0, C_1, C_2 in the singular integral $\mathscr{S}_x$ where $C_0=y_c-\overline{y}$, $C_1=R\,\cos\,\chi_i$, and $C_2=R\,\sin\,\chi_i$.

APPENDIX C: ELASTOACOUSTICS OF A BAFFLED FLAT STRIP BY MODAL ANALYSIS

Consider a baffled flat simply supported strip excited on its lower face by plane waves at an incident angle θ. Frequency response of the strip is governed by the equation

$$w''''-k^4w-2p/D=2P/D,$$

$$D=\frac{Eh^3}{12(1-\nu^2)},\quad k^4=\frac{\omega^2}{c_0^2 r_g^2}, \tag{C1}$$

$$c_0^2=\frac{E}{\rho(1-\nu^2)},\quad r_g^2=h^2/12,$$

where $(\,)'$ is derivative $w\cdot r\cdot t\cdot\ x$, and E, ρ, ν, h, l_x, ω are defined after Eq. (28) of the text, p is pressure radiated from the upper face, and P is incident pressure of the plane wave

$$P=P_0e^{ik(x\,\sin\,\theta+y\,\cos\,\theta)}. \tag{C2}$$

For simple supports

$$w(x)=\sum_j a_j\,\sin(k_{xj}x),\quad k_{xj}=j\pi/l_x. \tag{C3}$$

For a flat strip, the pressure $p(x)$ is expressed in terms of the 2-D Green's function and normal pressure gradient of the strip

$$p(x)=\int_0^{l_x}G_2(x-\xi)\,\frac{\partial p(\xi)}{\partial n}\,d\xi,$$

$$G_2(x)=-\tfrac{1}{2}\,[Y_0(k_f x)-iJ_0(k_f x)]. \tag{C4}$$

Continuity of acoustic and elastic velocities on the face of the strip yields

$$\frac{\partial p}{\partial n}=\rho_f\omega^2 w. \tag{C5}$$

Substituting (C2)–(C5) in (C1) and exploiting orthogonality of $\sin(k_{xj}x)$ produces equations in the coefficients a_m

$$a_m(k_{xm}^4-k^4)-\frac{2}{D}\sum_j P_{mj}a_j=\frac{2P_0}{D}\,\hat{p}_{0m},$$

$$P_{mj}=\rho_f\omega^2\int_0^{l_x}\int_0^{l_x}G_2(x-\xi)\sin(k_{xj}\xi)d\xi\,\sin(k_{xm}x)dx,$$

$$\hat{p}_{0m}=\begin{cases}\dfrac{2m\pi}{(m\pi)^2-(k_f l_x\,\sin\,\theta)^2}\,[1-(-1)^m e^{ik_f l_x\,\sin\,\theta}] \\ \quad m\pi\neq k_f l_x\,\sin\,\theta \\ i\quad m\pi=k_f l_x\,\sin\,\theta.\end{cases}\tag{C6}$$

Solving for a_m in (C6) determines $w(x)$ and $p(x)$. Substituting $p(x)$, $v^*(x)$, and $w(x)$ in expressions (26) and (27) for TL^θ and TL^D gives

$$\mathrm{TL}^\theta=\rho_f c_f\omega R_e\left\{\sum_m\sum_k ia_m P_{mk}a_k^*\right\}\Big/(P_0^2 l_x\cos\,\theta),\tag{C7}$$

$$\mathrm{TL}^D=\int_0^{\pi/2}(\mathrm{TL}^\theta)\cos\,\theta\,d\theta.$$

The integrable singularity in modal pressure has the form

$$p_{sj}(x)=\int_0^{l_x}\ln\,k_f(\xi-x)\sin(k_{xj}\xi)d\xi.\tag{C8a}$$

Integration by parts converts (C8a) to

$$p_{sj}(x)=\frac{1}{k_f}\ln(k_f/2)(\cos\,k_f l-1)+[\,\eta(\ln\,\eta-1)\times\sin(k_{xj}\xi)]_0^{l_x}-\int_0^{l_x}\eta(\ln\eta-1)\times k_{xj}\cos(k_{xj}\xi)d\xi,\quad \eta=\xi-x.\tag{C8b}$$

APPENDIX D: ACOUSTIC IMPEDANCE OF LINE SOURCE

Let (x,z) be an orthogonal coordinate system (see Fig. D1). Consider a finite line source of length l_x so that $-l_x/2\leqslant x\leqslant l_x/2$ and $z=0$. Let r_1 be the normal distance to a field point P_f at $(0,r_1)$. Assuming the line source density is constant and of unit strength, acoustic pressure p at P_f from the line source undergoing uniform periodic oscillation along z is

$$p(0,r_1)=2\rho_f c_f k_f\int_0^{l_x/2}\frac{e^{ik_f r}}{r}\,dx,\tag{D1}$$

$k_f=\omega/c_f$ and r is the distance between (x_0,z_0) and a point $(x,0)$ on the line source. Rewrite (D1) as

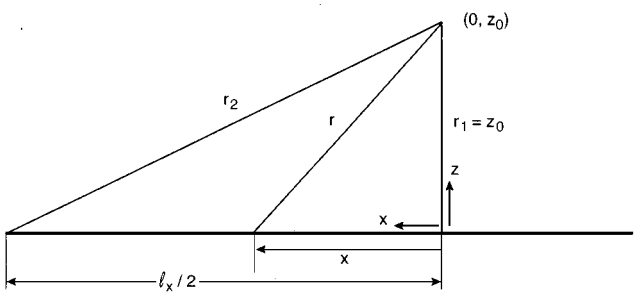


FIG. D1. Line source.

$$\frac{1}{\rho_f c_f}p(0,r_1)=\widetilde{Z}_{f0}=2k_f\int_0^{l_x/2}\frac{e^{ik_f r}}{r}\left(1-\frac{2x}{l_x}\right)dx+4k_f\int_0^{l_x/2}\frac{e^{ik_f r}}{r}\frac{x}{l_x}\,dx=2k_f\int_0^{l_x/2}\frac{e^{ik_f r}}{r}\left(1-\frac{2x}{l_x}\right)dx+\frac{4k_f}{l_x}\int_{r_1}^{r_2}e^{ik_f r}dr.\tag{D2}$$

The first term in (D2) can be approximated by

$$I_1=2k_f\int_0^{l_x/2}\frac{e^{ik_f r}}{r}\left(1-\frac{2x}{l_x}\right)dx\simeq\left(2.5\sqrt{\frac{k_f}{r_1}}-\frac{4}{l_x}\right)(\cos\,k_f r_1+i\,\sin\,k_f r_1).\tag{D3}$$

Exact evaluation of the second term in (D2) gives

$$I_2=\frac{4k_f}{l_x}\int_{r_1}^{r_2}e^{ik_f r}\,dr=\frac{4}{l_x}[\sin\,k_f r_2-\sin\,k_f r_1-i(\cos\,k_f r_2-\cos\,k_f r_1)].\tag{D4}$$

Taking the absolute value of $|I_1+I_2|$ yields

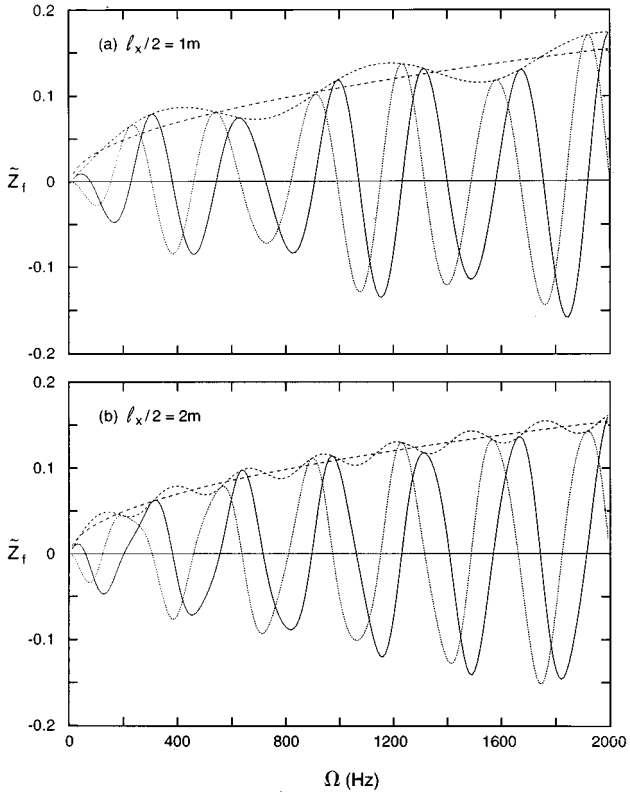


FIG. D2. Normalized point impedance $\widetilde{Z}_f$ from 3-D translating line source; ——— $R_e(\widetilde{Z}_f)$; ····· $I_m(\widetilde{Z}_f)$; -------- $|\widetilde{Z}_f|$; – – – – 2-D (a) $l_x/2=1$ m; (b) $l_x/2=2$ m; $r_1=1$ m.

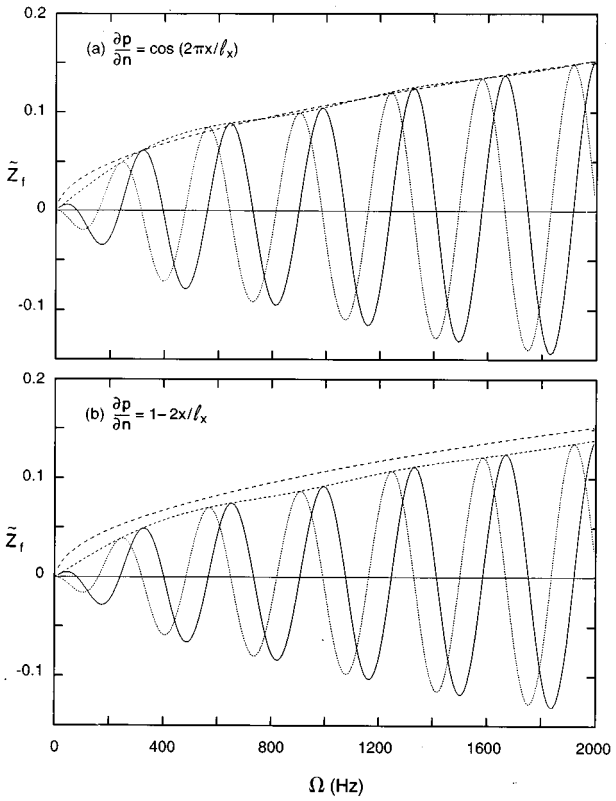


FIG. D3. Normalized point impedance $\tilde{Z}_f$ from 3-D line source; $l_x/2 = 1$ m, $r_1 = 1$ m. ——— $R_e(\tilde{Z}_f)$; ····· $I_m(\tilde{Z}_f)$; ------- $|\tilde{Z}_f|$; – – – – 2-D (a) $\partial p/\partial n = \cos(2\pi x/l_x)$, (b) $\partial p/\partial n = (1 - 2x/l_x)$.

$$|\tilde{Z}_{f0}| \simeq 2.5\sqrt{\frac{k_f}{r_1}} - \frac{4}{l_x}\cos(k_f \Delta r) + o\left(l_x^2\sqrt{\frac{k_f}{r_1}}\right)^{-1},$$

$$\Delta r = r_2 - r_1 = \sqrt{(l_x/2)^2 - r_1^2} - r_1, \tag{D5}$$

(D5) reveals that acoustic pressure from the line source is made of two parts: the first part, given by I_1, is smooth and varies with k_f like $2.5\sqrt{k_f/r_1}$, and the second part, given by I_2, fluctuates with k_f with amplitude $(4/l_x)$ and period

$$T_0 = \frac{1}{\tilde{\Omega}}, \quad \tilde{\Omega} = \frac{\Delta r}{c_f} \Rightarrow T_0 = \frac{c_f}{\Delta r}. \tag{D6}$$

Figure D2(a) and (b) plots $\tilde{Z}_{f0}$ evaluated numerically for $l_x/2 = 1$ m and 2 m. From (D6)

$$\frac{T_0(1m)}{T_0(2m)} = \frac{\sqrt{5}-1}{\sqrt{2}-1} \simeq 3.$$

This value agrees with the ratio of periods from Fig. D2(a) and (b). Figure D3(a) and (b) plots $\tilde{Z}_{f1}$ and I_1 versus frequency for the case in Fig. D2(a) where

$$\tilde{Z}_{f1} = 2k_f \int_0^{l_x/2} \frac{e^{ik_f r}}{r}\cos\left(\frac{2\pi x}{l_x}\right) dx. \tag{D7}$$

Note that the distribution $\cos(2\pi x/l_x)$ in (D7) reduces frequency and amplitude of fluctuation in Fig. D2(a). This also applies to I_1 where the distribution $(1 - 2x/l_x)$ in (D3) produces the same effect. Indeed, any piecewise continuous distribution vanishing at the ends $x = \pm l_x/2$ eliminates the fluctuation.

In the limit as $l_x \to \infty$, the integral in (D1) reduces to

$$\tilde{Z}_f = k_f \int_{-\infty}^{\infty} \frac{e^{ik_f r}}{r}\, dx = -i\pi k_f H_0^{(2)}(k_f r_1), \tag{D8}$$

which is identical to the 2-D impedance from a point source.

[1] J. E. Young, ''Transmission of sound through thin elastic plates,'' J. Acoust. Soc. Am. **26**, 485 (1954).
[2] E. C. Sewell, ''Transmission of reverberant sound through a single-leaf partition surrounded by an infinite rigid baffle,'' J. Sound Vib. **12**, 21–32 (1970).
[3] S. Alper and E. B. Magrab, ''Radiation from the forced harmonic vibrations of a clamped circular plate in an acoustic fluid,'' J. Acoust. Soc. Am. **48**, 681–691 (1970).
[4] S. N. Yousri and F. J. Fahy, ''Sound radiation from transversely vibrating unbaffled beams,'' J. Sound Vib. **26**, 437–439 (1973).
[5] G. H. Shmidt, ''Influence of an unbounded elastic plate in the radiation of sound by a point source,'' J. Sound Vib. **53**, 289–300 (1977).
[6] I. Nakayama, A. Nakamura, and R. Takeuchi, ''Sound insulation of a circular plate for a single sound pulse,'' Acustica **46**, 330–340 (1980).
[7] I. Nakayama, A. Nakamura, and R. Takeuchi, ''Sound insulation evaluated in radiation field of a vibrating circular plate,'' Acustica **47**, 89–96 (1981).
[8] R. W. Guy, ''The steady state transmission of sound at normal and oblique incidence through a thin panel backed by a rectangular room. A multi-modal analysis,'' Acustica **43**, 295–304 (1979).
[9] R. W. Guy, ''The transmission of airborne sound through a finite panel, air gap, panel and cavity configuration. A steady analysis,'' Acustica **49**, 323–333 (1981).
[10] S. Ljunggren, ''Airborne sound insulation of thin walls,'' J. Acoust. Soc. Am. **89**, 2324–2337 (1991).
[11] S. N. Beshenkov and T. D. Volkova, ''Forced vibration and acoustic properties of a rectangular plate that interacts with its environment,'' Mech. Solids **25**, 173–178 (1990).
[12] A. Berry, J. L. Guyader, and J. Nicholas, ''A general formulation for the sound radiation from rectangular, baffled plates with arbitrary boundary conditions,'' J. Acoust. Soc. Am. **88**, 2792–2802 (1990).
[13] A. Berry, ''A new formulation for vibrations and sound radiation of fluid-loaded plates with elastic boundary conditions,'' J. Acoust. Soc. Am. **96**, 889–901 (1994).
[14] D. Takahashi, ''Effects of panel boundness on sound transmission problems,'' J. Acoust. Soc. Am. **98**, 2598–2603 (1995).
[15] N. Atalla, J. Nicholas, and C. Gauthier, ''Acoustic radiation of an unbaffled plate with general elastic boundary conditions,'' J. Acoust. Soc. Am. **99**, 1484–1494 (1996).
[16] J. BenMariem and M. Hamdi, ''A new boundary finite element method for fluid-structure interaction,'' Int. J. Numer. Methods Eng. **24**, 1251–1267 (1987).
[17] F. Sgard and N. Atalla, ''Calcul du rayonnement acoustique d'une plaque non-bafflee par une methode variationelle d'elements frontieres,'' Internal report, GAUS, Universite de Sherbrooke, 1992a.
[18] M. A. Hamdi, ''Rayon User's Manual,'' STRACO SA, Compiegne, France, 1992.
[19] P. Guisset, ''SYSNOISE User's Manual,'' NIT N.V., Leuven, Belgium, 1992.
[20] F. Sgard and N. Atalla, ''GAUSVEM, User's Manual,'' Version 1, GAUS, Universite de Sherbrooke, 1992b.
[21] N. Vlahopoulos *et al.*, ''COMET/ACOUSTICS user document,'' Automated Analysis, Inc., July 1994.
[22] M. Verkamen, ''Laboratory noise measurements on polycarbonate panels,'' Peutz and Associates consulting firm B.V., The Netherlands, Report No. A-456-1, 1993].
[23] E. L. Ince, *Ordinary Differential Equations* (Dover, New York, 1953), 2nd ed.

Closed form solutions to the vibroacoustic response of strip structures

Carina Ting
Cambridge Acoustical Associates, Medford, Massachusetts 02155

I. Y. Shen
Department of Mechanical Engineering, Box 352600, University of Washington, Seattle, Washington 98195-2600

(Received 9 August 1996; accepted for publication 24 April 1997)

This paper presents some closed form expressions for the radiated power from flat elastic radiating strip structures of infinite length and finite width in an ambient light acoustic medium. New analytical expressions for the structural displacements, the wave number transforms of the structural displacements, the far-field pressure, and the total radiated sound power associated with an arbitrary strip that has been subjected to constant frequency loads including concentrated loads, boundary loads, or incident plane pressure waves are developed through the method of distributed transfer functions. With the exception of the radiated power from a structure subjected to a plane pressure wave, all of the expressions obtained are in closed form. The solutions presented are independent of the specific differential equations and boundary conditions governing the structure. An example analysis of a structure consisting of a bare aluminum strip with an active control scheme using boundary actuation that is subjected to an incident plane pressure wave is performed to illustrate the advantages of the new solutions. © *1997 Acoustical Society of America.*
[S0001-4966(97)02908-1]

PACS numbers: 43.40.Dx, 43.40.Rj [CBB]

INTRODUCTION AND MOTIVATION

The objective of this paper is to offer a new analysis tool for the evaluation and understanding of noise and vibration treatments applied to flat plates vibrating in a light ambient fluid. In particular, this paper presents closed form solutions to the classic radiating strip corresponding to three specific excitations: loads applied to a boundary, a concentrated load, and a traveling pressure wave. Since the ambient fluid is light, fluid–structure interaction mechanisms will not be considered in this paper.

The procedure for determining the radiated acoustic pressure from a flat structure vibrating in a light fluid is straightforward in principle. First, the normal velocities of the structure are determined, ignoring fluid loading. Then the resulting radiated pressure can be determined from the vibrating interface into the light fluid using linear acoustic theory.[1] The radiated pressure will typically have a directivity pattern, complicating the issue of comparing the sound reduction achieved from one structure to another. A better metric of radiated sound reduction is usually the total radiated sound power. Assuming steady state harmonic excitations, the average sound power flowing from the structure can be determined by integrating the sound intensity. Equivalently, in this simple case, the average radiated sound power can also be determined by the integration of the wave number transform of the normal velocity of the flat structure.[1] This sort of procedure applied to one-dimensional structures is well demonstrated in papers by Guigou and Fuller[2] and Baumann *et al.*[3] for the purpose of active noise control. The procedure to be described in this paper differs from those presented in Refs. 2 and 3, in that the structure's total radiated power is obtained in closed form, as opposed to the usual resort to numerical integration to obtain this quantity. This is achieved through use of the method of distributed transfer functions developed by Yang and Tan,[4] through which the structural displacement is expressed in a form that allows the closed form integration of the total radiated power expression.[5]

The solutions presented here offer many advantages. First, most of the expressions are given in closed form, avoiding tedious numerical integrations. Second, the closed form expressions facilitate the application of optimization techniques, because the derivatives of the objective functions with respect to design variables can also be obtained in closed form. Finally, the solutions are independent of the specific differential equations and boundary conditions that may describe the strip structure. For example, multi-layer plates are easily treated by this method.[6] Also, the solutions are in a form that has relatively light computational requirements and variations of some of the parameters describing a structure, including the boundary conditions, may be computed with little additional effort.

In this paper, the closed form analytic solution is presented for a general, flat, one-dimensional continuum in an operator form. In particular, specific solutions for the following three excitations will be described: boundary loads, concentrated loads, and traveling pressure waves. Verifications and demonstration of the new solutions were performed on an example case of a bare elastic plate/strip. For verification, the analytical results are compared with those independently obtained from the numerical evaluation of the Kirchhoff–Helmholtz integral equation through a finite element analysis. To demonstrate the application of the closed form ex-

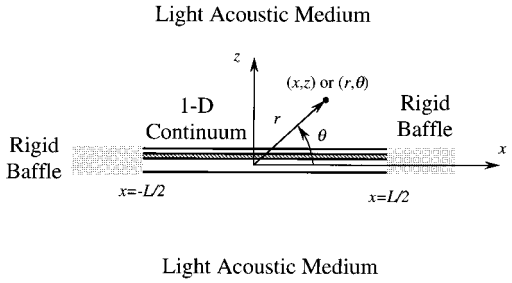


FIG. 1. The geometry and coordinate system of the radiating strip structure.

pression to active noise control, a boundary control scheme based on Lyapunov's direct method is proposed to reduce the total radiated power of the bare aluminum plate subjected to an incident plane pressure wave. Lastly, the advantages and disadvantages of these techniques that the authors have discovered from applying them are discussed.

I. GENERAL SOLUTIONS FOR AN ARBITRARY EXCITATION

The types of structures to be analyzed here are flat, one-dimensional structures for which the differential equations of motion can be cast in the form

$$\mathbf{L}[\mathbf{u}(x,t)]=\mathbf{f}(x,t), \tag{1}$$

where $\mathbf{L}[\cdot]$ is a linear, time invariant, differential operator in matrix form and $\mathbf{u}$ is a vector varying with x, one dimension in space, and t, time. The vector $\mathbf{u}$ must include the displacement normal to the flat surface and the vector $\mathbf{f}$ includes the forcing term. These types of structures are often referred to as radiating strips in many textbooks. Figure 1 describes the geometry of the elastic structure assumed in Eq. (1) as well as a coordinate system used for the subsequent discussion.

A total width of L is assigned, with the structure defined such that $-L/2\leqslant x\leqslant +L/2$. The region $|x|>L/2$ is assumed to have zero normal velocity and is rigid in the acoustic sense. Each half-space $z>0$ and $z<0$ consists of an acoustic fluid fully described by the properties ρ_0, the mean mass density of the fluid, and c, the compressional wave speed of the fluid. It is assumed in this paper that the ambient fluid is light relative to the structure and fluid loading will be ignored.

The boundary conditions necessary for a unique solution of Eq. (1) can also be described compactly in the following linear operator form:

$$\mathbf{Q}[\mathbf{u}(x=-L/2,t),\ \mathbf{u}(x=+L/2,t)]=\boldsymbol{\epsilon}(t), \tag{2}$$

where $\mathbf{Q}[\cdot]$ is a matrix linear differential operator operating on the vectors $\mathbf{u}(x=-L/2,t)$ and $\mathbf{u}(x=+L/2,t)$. The term $\boldsymbol{\epsilon}(t)$ allows for nonhomogenous boundary conditions. At this point, it is convenient to assume displacements and forcing functions to consist of a steady state harmonic time variation, i.e.,

$$\mathbf{u}(x,t)=\mathrm{Re}\{\hat{\mathbf{u}}(x,\omega)e^{+i\omega t}\}, \tag{3a}$$

$$\mathbf{f}(x,t)=\mathrm{Re}\{\hat{\mathbf{f}}(x,\omega)e^{+i\omega t}\}, \tag{3b}$$

$$\boldsymbol{\epsilon}(t)=\mathrm{Re}\{\hat{\boldsymbol{\epsilon}}(\omega)e^{+i\omega t}\}. \tag{3c}$$

In the following equations, quantities with $\hat{}$ represent the steady state components.

The first step in the procedure is to recast Eq. (1), given Eqs. (3a)–(3c), into the state space form in x

$$\frac{\partial}{\partial x}[\hat{\mathbf{y}}(x,\omega)]=\mathbf{F}(\omega)\cdot\hat{\mathbf{y}}(x,\omega)+\hat{\mathbf{q}}(\omega) \tag{4}$$

where the vector $\hat{\mathbf{y}}$ is an $N\times 1$ vector consisting of space derivatives of the elements of $\hat{\mathbf{u}}$, $\mathbf{F}$ is an $N\times N$ matrix describing the characteristic dynamic behavior of the structure, and $\hat{\mathbf{q}}$ is an $N\times 1$ vector consisting of the elements of $\hat{\mathbf{f}}$ and zeros. The number N, in general, is determined by the total order of derivative of $\mathbf{L}[\cdot]$ in (1). Similarly, boundary condition (2) is recast into the following state space form:

$$\mathbf{M}(\omega)\cdot\hat{\mathbf{y}}(x=-L/2,\omega)+\mathbf{N}(\omega)\cdot\hat{\mathbf{y}}(x=+L/2,\omega)=\hat{\mathbf{g}}(\omega), \tag{5}$$

where $\mathbf{M}$ and $\mathbf{N}$ are $N\times N$ matrices and $\hat{\mathbf{g}}$ is an $N\times 1$ vector consisting of elements of $\hat{\boldsymbol{\epsilon}}$ and zeros.

According to Yang and Tan,[4] the solution to Eqs. (4) and (5) is

$$\hat{\mathbf{y}}(x,\omega)=\int_{-L/2}^{+L/2}\tilde{\mathbf{G}}(x,\xi,\omega)\cdot\hat{\mathbf{q}}(\xi,\omega)\,d\xi+\mathbf{H}(x,\omega)\cdot\hat{\mathbf{g}}(x,\omega), \tag{6}$$

where

$$\tilde{\mathbf{G}}(x,\xi,\omega)=\begin{cases}+\mathbf{e}^{\mathbf{F}x}\cdot[\mathbf{M}\mathbf{e}^{-\mathbf{F}L/2}+\mathbf{N}\mathbf{e}^{\mathbf{F}L/2}]^{-1}\mathbf{M}\mathbf{e}^{\mathbf{F}(-L/2-\xi)}, & \xi<x,\\ -\mathbf{e}^{\mathbf{F}x}\cdot[\mathbf{M}\mathbf{e}^{\mathbf{F}L/2}+\mathbf{N}\mathbf{e}^{+\mathbf{F}L/2}]^{-1}\mathbf{N}\mathbf{e}^{\mathbf{F}(+L/2-\xi)}, & \xi>x,\end{cases} \tag{7}$$

and

$$\mathbf{H}(x,\omega)=\mathbf{e}^{\mathbf{F}x}\cdot[\mathbf{M}\mathbf{e}^{-\mathbf{F}L/2}+\mathbf{N}\mathbf{e}^{+\mathbf{F}L/2}]^{-1}. \tag{8}$$

In (7) and (8), the matrix exponential $\mathbf{e}^{\mathbf{F}x}$ can be determined by the expression

$$\mathbf{e}^{\mathbf{F}x}=\mathbf{V}e^{\mathbf{\Lambda}x}\mathbf{V}^{-1}, \tag{9}$$

where $\mathbf{V}$ is the modal matrix and $\mathbf{\Lambda}$ is the diagonal matrix consisting of the eigenvalues λ_m of the $\mathbf{F}$ matrix, i.e.,

$$\Lambda=\mathrm{diag}[\lambda_1,\lambda_2,\ldots,\lambda_N]. \tag{10}$$

Note that this expression for the matrix exponential is valid only when each eigenvalue of $\mathbf{F}$ is distinct. Furthermore, the closed form solution to be derived later in this paper requires that no eigenvalue of $\mathbf{F}$ be purely imaginary, which is easily accomplished by including a loss factor into the stiffness parameters associated with the structure. The reason for this requirement will become more apparent later when the closed form solution is derived.

Let the βth component of $\hat{\mathbf{y}}$, denoted by $\hat{y}_\beta$, be the displacement $\hat{w}(x,\omega)$ normal to the strip structure. Then the steady state normal velocity $\hat{\dot{w}}(x,\omega)$ of the interface to the ambient medium to which the sound will radiate is

$$\hat{\dot{w}}(x,\omega)=i\omega\hat{w}(x,\omega)=i\omega\hat{y}_\beta, \tag{11}$$

where the steady state assumption has been used. Moreover, Cremer[1] gives the following expression for the total radiated sound power:

$$\langle \mathbf{P}(\omega)\rangle = \frac{\rho_0 ck}{4\pi}\int_{-k}^{+k}\frac{|\tilde{\hat{w}}(k_x,\omega)|^2}{\sqrt{k^2-k_x^2}}\,dk_x, \tag{12}$$

where the function $\tilde{\hat{w}}(k_x,\omega)$ in Eq. (12) is the wave number transform of the normal velocity $\hat{w}(x,\omega)$ defined by

$$\tilde{\hat{w}}(k_x,\omega)=\int_{-L/2}^{+L/2}\hat{w}(x,\omega)e^{+ik_x x}\,dx. \tag{13}$$

(In the following equations, quantities with $\tilde{}$ denote the wave number transform.) Given an expression for the wave number transform, the far-field radiated pressure is obtained by the following expression[7] that uses the polar coordinates indicated in Fig. 1:

$$\hat{p}(r,\theta,\omega)\approx\rho_0\omega\frac{e^{-i(kr-\pi/4)}}{\sqrt{2\pi kr}}\tilde{\hat{w}}(k_x=k\cos\theta,\omega),$$
$$kr\gg 1,\quad r\gg L, \tag{14}$$

where k is the acoustic wave number ω/c.

Using Eqs. (6)–(14), closed form expressions of the radiated power for specific excitations can be obtained in terms of $\hat{\mathbf{f}}(x)$ and $\hat{\boldsymbol{\epsilon}}(x)$. Three such cases will follow.

II. BOUNDARY LOAD

Referring to Eqs. (1)–(3), boundary load is applied to the structure by setting $\hat{\mathbf{f}}(x)=\mathbf{0}$ and $\hat{\boldsymbol{\epsilon}}=\hat{\boldsymbol{\epsilon}}(\omega)$. Therefore, (6), (8), and (9) lead to

$$\hat{\mathbf{y}}(x,\omega)=\mathbf{V}\mathbf{e}^{\boldsymbol{\Lambda}x}\overline{\boldsymbol{\psi}}(\omega), \tag{15}$$

where the vector $\overline{\boldsymbol{\psi}}(\omega)$ is defined by

$$\overline{\boldsymbol{\psi}}(\omega)\equiv[\mathbf{M}(\omega)\mathbf{V}\mathbf{e}^{-\boldsymbol{\Lambda}L/2}+\mathbf{N}(\omega)\mathbf{V}\mathbf{e}^{+\boldsymbol{\Lambda}L/2}]^{-1}\hat{\mathbf{g}}(\omega). \tag{16}$$

According to (11) and (15),

$$\hat{w}(x,\omega)=\sum_{m=1}^{N}\psi_m e^{\lambda_m x}, \tag{17}$$

where $\lambda_m=\Lambda_{mm}$ is the mth eigenvalue of $\mathbf{F}$ and ψ_m is

$$\psi_m\equiv i\omega\nu_{\beta m}\overline{\psi_m}, \tag{18}$$

with $\nu_{\beta m}$ and $\overline{\psi_m}$ being the elements of $\mathbf{V}$ and $\overline{\boldsymbol{\psi}}(\omega)$, respectively. Substitution of (17) into (13) gives the wave number transform

$$\tilde{\hat{w}}(k_x,\omega)=\sum_{m=1}^{N}\psi_m\frac{e^{(\lambda_m+ik_x)L/2}-e^{-(\lambda_m+ik_x)L/2}}{\lambda_m+ik_x}. \tag{19}$$

Also, substitution of (17) into (14) gives the approximate far-field radiated pressure

$$\hat{p}(r,\theta,\omega)\approx\rho_0\omega\frac{e^{-i(kr-\pi/4)}}{\sqrt{2\pi kr}}\cdot\sum_{m=1}^{N}\psi_m\frac{e^{(\lambda_m+ik\cos\theta)L/2}-e^{-(\lambda_m+ik\cos\theta)L/2}}{\lambda_m+ik\cos\theta},$$
$$kr\gg 1,\quad r\gg L. \tag{20}$$

Substitution of (19) into (12) gives the radiated power

$$\langle \mathbf{P}(\omega)\rangle = \frac{\rho_0 ck}{4\pi}\sum_{m=1}^{N}\sum_{n=1}^{N}\psi_m^*\psi_n\{[e^{(\lambda_m^*+\lambda_n)L/2}+e^{-(\lambda_m^*+\lambda_n)L/2}]I_{mn}(0)-e^{(\lambda_m^*-\lambda_n)L/2}I_{mn}(-L)-e^{-(\lambda_m^*-\lambda_n)L/2}I_{mn}(L)\}, \tag{21}$$

where the * superscript is the complex conjugate and $I_{mn}(\cdot)$ is an integral defined as

$$I_{mn}(\delta_0)\equiv\int_{-k}^{+k}\frac{e^{ik_x\delta_0}}{(\lambda_m^*-ik_x)(\lambda_n+ik_x)\sqrt{k^2-k_x^2}}\,dk_x. \tag{22}$$

According to Appendix A, $I_{mn}(\delta_0)$ can be further recast into the following contour integral:

$$I_{mn}(\delta_0)=\frac{2i}{k^2}\oint_{|z|=1}\frac{z\cdot[J_0(k\delta_0)+2\cdot\Sigma_{\nu=1}^{\infty}J_\nu(k\delta_0)\cdot z^\nu]}{(z-P_m^{(1)})(z-P_m^{(2)})(z-P_n^{(1)})(z-P_n^{(2)})}\,dz, \tag{23}$$

where $J_\nu(\cdot)$ is the Bessel function of the first kind and $P_m^{(1)}$, $P_m^{(2)}$, $P_n^{(1)}$, and $P_n^{(2)}$ are poles defined by

$$P_m^{(1,2)}=+\frac{\lambda_m^*}{k}\pm\sqrt{\left(\frac{\lambda_m^*}{k}\right)^2+1}, \tag{24}$$

$$P_n^{(1,2)}=-\frac{\lambda_n}{k}\pm\sqrt{\left(\frac{\lambda_n}{k}\right)^2+1}. \tag{25}$$

The contour integral (23) is then evaluated as $2\pi i$ times the sum of the residues given the poles $P_m^{(1)}$, $P_m^{(2)}$, $P_n^{(1)}$, and $P_n^{(2)}$ that lie within the unit circle $|z|=1$ on the complex plane. Note that $|P_m^{(1)}P_m^{(2)}|=1$ and $|P_n^{(1)}P_n^{(2)}|=1$. Therefore, $P_m^{(1)}$ and $P_m^{(2)}$ will not be inside the unit circle at the same time, and neither will $P_n^{(1)}$ and $P_n^{(2)}$. Besides, the expression in the contour integral in (23) converges uniformly at each residue evaluation.

Substitution of (23) into (21) leads to a closed form matrix expression for the radiated power per unit length of the strip

$$\langle \mathbf{P}(\omega)\rangle = \frac{\rho_0 ck}{4\pi}\boldsymbol{\Psi}^H\boldsymbol{\Xi}\boldsymbol{\Psi}, \tag{26}$$

where $\boldsymbol{\Psi}$ is an $N\times 1$ column vector consisting of the coefficients ψ_m, and $\boldsymbol{\Psi}^H$ is the Hermitian (the complex conjugate and transpose) of $\boldsymbol{\Psi}$. The matrix $\boldsymbol{\Xi}$ is an $N\times N$ matrix consisting of the elements ξ_{mn} given by

$$\xi_{mn}=\frac{2i}{k^2}\oint_{|z|=1}\frac{z\{A+(B+C)J_0(kL)+2\Sigma_{\nu=1}^{\infty}[B(-1)^{\nu}+C]J_{\nu}(kL)z^{\nu}\}}{(z-P_m^{(1)})(z-P_m^{(2)})(z-P_n^{(1)})(z-P_n^{(2)})},\tag{27}$$

where

$$A=e^{(\lambda_m^*+\lambda_n)L/2}+e^{-(\lambda_m^*+\lambda_n)L/2},$$

$$B=-e^{(\lambda_m^*-\lambda_n)L/2},\quad C=-e^{-(\lambda_m^*-\lambda_n)L/2}.$$

Note that the Bessel function series evaluations may be shared between each evaluation of each entry of the matrix ξ_{mn}. In addition, the Ξ matrix is Hermitian; therefore, only its upper diagonal needs to be calculated explicitly.

As a final remark, the closed form formulation is not analytic when **F** has pure imaginary eigenvalues. In this case, the denominator of (22) vanishes and the poles of (23) lie on the unit circle, which is the integration contour. A separate treatment is needed to handle this special case.

III. POINT DRIVE

Referring to Eqs. (1)–(3), a concentrated load is applied to the structure with homogenous boundary conditions by setting $\hat{\mathbf{f}}(x,\omega)=\hat{\mathbf{f}}_0(\omega)\cdot\delta(x-x_0)$ and $\hat{\boldsymbol{\epsilon}}(\omega)=\mathbf{0}$. The vector $\hat{\mathbf{f}}_0$ is a spatially independent vector and $\delta(\cdot)$ indicates the Dirac delta function.

Performing the first step, one sets $\hat{\mathbf{q}}(x)$ in Eq. (4) to be

$$\hat{\mathbf{q}}(x)=\boldsymbol{\gamma}\delta(x-x_0),\tag{28}$$

where $\boldsymbol{\gamma}$ is a spatially independent vector such that $\hat{\mathbf{f}}(x)$ and Eqs. (1), (3), and (4) are consistent. Evaluation of Eq. (6) with (7) and (9) yields

$$\hat{\mathbf{y}}(x,\omega)=\begin{cases}-\mathbf{V}(\omega)\mathbf{e}^{\boldsymbol{\Lambda}(\omega)x}\cdot\overline{\boldsymbol{\sigma}}(\omega), & x<x_0,\\ +\mathbf{V}(\omega)\mathbf{e}^{\boldsymbol{\Lambda}(\omega)x}\cdot\overline{\boldsymbol{\tau}}(\boldsymbol{\omega}), & x>x_0,\end{cases}\tag{29}$$

where

$$\overline{\boldsymbol{\sigma}}\equiv[\mathbf{MV}\mathbf{e}^{-\boldsymbol{\Lambda}\cdot L/2}+\mathbf{NV}\mathbf{e}^{\boldsymbol{\Lambda}\cdot L/2}]^{-1}\mathbf{NV}\mathbf{e}^{\boldsymbol{\Lambda}(L/2-x_0)}\mathbf{V}^{-1}\boldsymbol{\gamma},\tag{30}$$

$$\overline{\boldsymbol{\tau}}\equiv[\mathbf{MV}\mathbf{e}^{-\boldsymbol{\Lambda}\cdot L/2}+\mathbf{NV}\mathbf{e}^{\boldsymbol{\Lambda}\cdot L/2}]^{-1}\mathbf{MV}\mathbf{e}^{\boldsymbol{\Lambda}(-L/2-x_0)}\mathbf{V}^{-1}\boldsymbol{\gamma}.\tag{31}$$

The normal velocity at the interface between the structure and the acoustic half-space is

$$\hat{w}(x,\omega)=\begin{cases}-\sum_m^N\ \sigma_m e^{\lambda_m x}, & x\leqslant x_0,\\ +\sum_m^N\ \tau_m e^{\lambda_m x}, & x\geqslant x_0,\end{cases}\tag{32}$$

where

$$\sigma_m\equiv i\omega\cdot\overline{\sigma_m}\nu_{\beta m},\tag{33}$$

$$\tau_m\equiv i\omega\cdot\overline{\tau_m}\nu_{\beta m}.\tag{34}$$

As in the case of the boundary load treated previously, expressions for the wave number transform of $\hat{w}(x,\omega)$ and the far-field radiated pressure are easily obtained by substitution of Eq. (32) into Eqs. (13) and (14). They are

$$\tilde{w}(k_x,\omega)=\sum_{m=1}^{N}\frac{\sigma_m e^{-(\lambda_m+ik_x)L/2}-(\sigma_m+\tau_m)e^{(\lambda_m+ik_x)x_0}+\tau_m e^{(\lambda_m+ik_x)L/2}}{\lambda_m+ik_x}$$

and

$$\hat{p}(r,\theta,\omega)\approx\rho_0\omega\cdot\frac{e^{-i(kr-\pi/4)}}{\sqrt{2\pi kr}}\cdot\sum_{m=1}^{N}\frac{\sigma_m e^{-(\lambda_m+ik\cos\theta)L/2}-(\sigma_m+\tau_m)e^{-(\lambda_m+ik\cos\theta)x_0}+\tau_m e^{(\lambda_m+ik\cos\theta)L/2}}{\lambda_m+ik\cos\theta},\quad kr\gg 1,\quad r\gg L.$$

The closed form evaluation of the total radiated power via Eq. (12) can be expressed as

$$\langle\mathbf{P}(\omega)\rangle=\frac{\rho ck}{4\pi}\{\mathbf{a}^H\mathbf{I}(0)\mathbf{a}+\mathbf{b}^H\mathbf{I}(0)\mathbf{b}+\mathbf{c}^H\mathbf{I}(0)\mathbf{c}+2\,\mathrm{Re}[-\mathbf{a}^H\mathbf{I}(L/2+x_0)\mathbf{b}+\mathbf{a}^H\mathbf{I}(L)\mathbf{c}-\mathbf{b}^H\mathbf{I}(L/2-x_0)\mathbf{c}]\},\tag{35}$$

where

$$\mathbf{a}=\mathbf{e}^{-\boldsymbol{\Lambda}L/2}\cdot\boldsymbol{\sigma},\tag{36}$$

$$\mathbf{b}=\mathbf{e}^{+\boldsymbol{\Lambda}x_0}\cdot(\boldsymbol{\sigma}+\boldsymbol{\tau}),\tag{37}$$

$$\mathbf{c}=\mathbf{e}^{\boldsymbol{\Lambda}L/2}\cdot\boldsymbol{\tau}.\tag{38}$$

In (35), $\mathbf{I}(\cdot)$ is a square matrix whose element $I_{mn}(\cdot)$ is given by (23). In addition, $\boldsymbol{\sigma}$ and $\boldsymbol{\tau}$ are $N\times 1$ column vectors with elements σ_m and τ_m given by (33) and (34), respectively. Also note that $\mathbf{I}(\cdot)$ does not have any symmetry as opposed to Ξ in the boundary load case. Nevertheless, the Bessel function evaluations are the same for each entry.

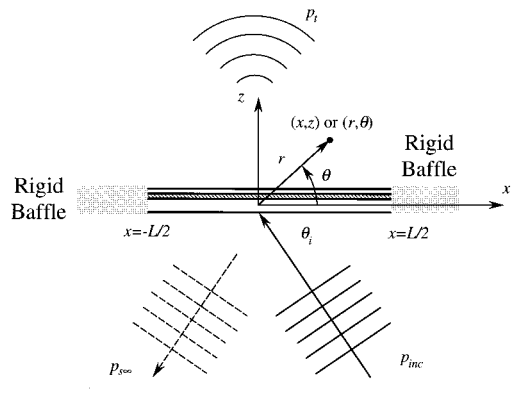


FIG. 2. The geometry and coordinate system of the radiating strip subjected to an impinging plane pressure wave.

IV. TRAVELING PRESSURE WAVE

Figure 2 shows an incident pressure wave originating from the $z<0$ half-space described by $p_{\text{inc}}(x,z,t,\theta_i) = p_i \exp[i(\omega t + kx \cos\theta_i + kz \sin\theta_i)]$ with $-\pi < \theta_i < 0$, where p_i is a constant amplitude coefficient and θ_i is the incidence angle. As a result of the elasticity of the strip, the pressure wave will be partly reflected and partly transmitted. Because negligible fluid loading is assumed, the transmitted pressure load p_t acting on the strip structure will be neglected and the reflected pressure load will be approximated by

$$p_{s\infty}(x,z,t,\theta_i) = p_i \exp[i(\omega t + kx \cos\theta_i - kz \sin\theta_i)], \tag{39}$$

which is the wave function that would be scattered from a rigid boundary.[7] Therefore, the loading acting on the strip structure is the sum of p_{inc} and $p_{s\infty}$ resulting in

$$\frac{\partial}{\partial x}[\hat{\mathbf{y}}(x,\omega)] = \mathbf{F}(\omega)\cdot\hat{\mathbf{y}}(x,\omega) + 2p_i \exp(ikx\cos\theta_i)\boldsymbol{\gamma} \tag{40}$$

and homogenous boundary conditions $\hat{\boldsymbol{\epsilon}}(\omega) = \mathbf{0}$. Through the same procedure, the resulting displacement vector becomes

$$\hat{\mathbf{y}}(x,\omega) = p_i\mathbf{V}\{e^{+ikx\cos\theta_i}\cdot\overline{\boldsymbol{\kappa}} - \mathbf{e}^{\boldsymbol{\Lambda}x}\cdot\overline{\boldsymbol{\phi}}\}. \tag{41}$$

In (41),

$$\overline{\boldsymbol{\kappa}}(\omega,\theta_i) \equiv 2[(ik\cos\theta_i)\mathbf{I}_N - \boldsymbol{\Lambda}]^{-1}\mathbf{V}^{-1}\boldsymbol{\gamma} \tag{42}$$

and

$$\overline{\boldsymbol{\phi}}(\omega,\theta_i) \equiv [\mathbf{MV}\mathbf{e}^{-\boldsymbol{\Lambda}L/2} + \mathbf{NV}\mathbf{e}^{+\boldsymbol{\Lambda}L/2}]^{-1} \times\{\mathbf{MV}e^{-(ik\cos\theta_i)L/2} + \mathbf{NV}e^{+(ik\cos\theta_i)L/2}\}\overline{\boldsymbol{\kappa}}, \tag{43}$$

where $\mathbf{I}_N$ is an $N\times N$ identity matrix. As in the previous load cases, the appropriate normal velocity of the strip is easily obtained from the displacement vector with

$$\hat{w}(x,\omega) = p_i\left\{\boldsymbol{\kappa}\cdot e^{+ikx\cos\theta_i} - \sum_{m=1}^{N}\phi_m e^{\lambda_m x}\right\}, \tag{44}$$

where

$$\boldsymbol{\kappa} = i\omega\sum_{m=1}^{N}\nu_{\beta m}\overline{\kappa_m} \tag{45}$$

and

$$\phi_m = i\omega\nu_{\beta m}\overline{\phi_m}. \tag{46}$$

The function $\boldsymbol{\kappa}$ is a characteristic of the system and depends on the driving frequency ω and the incidence angle θ_i. Similarly, ϕ_m is also a system characteristic and depends on the boundary conditions in addition to ω and θ_i.

Expressions for the wave number transform of $\hat{w}(x,\omega)$ and the far-field transmitted pressure are obtained by substitution of Eq. (44) into Eqs. (13) and (14). An expression for the wave number transform is

$$\tilde{\hat{w}}_t(k_x,\omega) = \kappa L\,\frac{\sin[(k_x + k\cos\theta_i)L/2]}{(k_x + k\cos\theta_i)L/2} - \sum_{m=1}^{N}\phi_m L\,\frac{\sin[(k_x - i\lambda_m)L/2]}{(k_x - i\lambda_m)L/2}. \tag{47}$$

An expression for the far-field radiated pressure is

$$\hat{p}_t(r,\theta,\omega) \approx \rho_0\omega\cdot\frac{e^{-i(kr-\pi/4)}}{\sqrt{2\pi kr}}\cdot\left\{\kappa L\,\frac{\sin[(\cos\theta + \cos\theta_i)kL/2]}{(\cos\theta + \cos\theta_i)kL/2} - \sum_{m=1}^{N}\phi_m L\,\frac{\sin[(\cos\theta - i\lambda_m/k)kL/2]}{(\cos\theta - i\lambda_m/k)kL/2}\right\}. \tag{48}$$

An expression for the transmitted power is

$$\langle\mathbf{P}(\omega,\theta_i)\rangle = \frac{\rho_0 ck}{4\pi}\{K(\omega,\theta_i) - 2\,\mathrm{Re}(\boldsymbol{\kappa}^*\boldsymbol{\Omega}\boldsymbol{\phi}) + \boldsymbol{\phi}^H\boldsymbol{\Xi}\boldsymbol{\phi}\}, \tag{49}$$

where $\boldsymbol{\Xi}$ is an $N\times N$ matrix whose elements are given in (27) and

$$K(\omega,\theta_i) = |\kappa|^2L^2\int_0^{\pi}\frac{\sin^2[kL(\cos\theta + \cos\theta_i)/2]}{[kL(\cos\theta + \cos\theta_i)/2]^2}\,d\theta \tag{50}$$

and

$$\Omega_m(\omega,\theta_i) = L^2\int_0^{\pi}\frac{\sin[kL(\cos\theta + \cos\theta_i)/2]}{kL(\cos\theta + \cos\theta_i)/2}\cdot\frac{\sin[kL(\cos\theta - i\lambda_m/k)/2]}{kL(\cos\theta - i\lambda_m/k)/2}\,d\theta. \tag{51}$$

Apparently the terms $K(\omega,\theta_i)$ and $\boldsymbol{\Omega}$ in (50) and (51) cannot be found in closed form via the straightforward evaluation of the contour integral. Therefore, $K(\omega,\theta_i)$ and $\boldsymbol{\Omega}$ need to be integrated numerically. Fortunately, the integrands in (50) and (51) are finite over the entire range of integration. Also note that part of the radiated power does not depend on the boundary conditions, as is described by $K(\omega,\theta_i)$ in (49).

V. EXAMPLE: BARE ALUMINUM PLATE

To illustrate the use and advantages of the methods presented here, an example analysis of a simply supported bare

thin aluminum plate subjected to an harmonic acoustic loading $2p_i \exp(i\omega t + ikx \cos\theta_i)$ is performed. This acoustic loading would model an incident plane pressure wave with amplitude p_i and incidence angle θ_i, provided that the pressure of the transmitted acoustic wave is much smaller than that of the incident wave. The numerical example consists of two parts. The first part is to compare the analytic solutions with a finite element solution in order to demonstrate the accuracy of the analytic solutions. The second part is to apply the analytic solutions together with a boundary active control scheme to reduce acoustic radiation.

The structural dynamics of a thin plate subjected to the harmonic loading is governed by

$$\left(D\frac{\partial^4}{\partial x^4}+\rho_p h_p\frac{\partial^2}{\partial t^2}\right)w(x,t) = 2p_i \exp(i\omega t + ikx\cos\theta_i), \tag{52}$$

where ρ_p is the mass density per volume of the plate material, h_p is the thickness of the plate, and D is the flexural rigidity of the plate given by

$$D=\frac{E_p(1+i\eta_p)h_p^3}{12(1-\nu_p^2)}. \tag{53}$$

In (53), ν_p is the Poisson's ratio, E_p is the Young's modulus, and η_p is the loss factor. Note that the loss factor η_p is necessary to ensure that the system has no purely imaginary eigenvalues as assumed in this paper.

The boundary conditions of the plate with or without the boundary control can be described by the following equations simultaneously. They are

$$w(x=-L/2,t)=\frac{\partial^2}{\partial x^2}w(x=-L/2,t) = w(x=+L/2,t)=0 \tag{54}$$

and a boundary moment $\mathbf{M}(t)$ actuated via the following control law

$$\mathbf{M}(t)=D\frac{\partial^2}{\partial x^2}w(x=+L/2,t) = -D\mu\frac{\partial^2}{\partial x\,\partial t}w(x=+L/2,t), \tag{55}$$

where μ is the control gain. When $\mu=0$, the boundary conditions (54) and (55) indicate that the plate is simply supported. When $\mu>0$, the boundary control scheme is asymptotically stable as shown in Appendix B through the Lyapunov's direct method.[8] To cast Eq. (52) into the form described in Eq. (4), define

$$\hat{\mathbf{y}}(x,\omega)=\left[\hat{w}\quad\frac{d\hat{w}}{dx}\quad\frac{d^2\hat{w}}{dx^2}\quad\frac{d^3\hat{w}}{dx^3}\right]^T, \tag{56}$$

$$\mathbf{F}(\omega)=\begin{bmatrix}0&1&0&0\\0&0&1&0\\0&0&0&1\\k_f^4&0&0&0\end{bmatrix}, \tag{57}$$

TABLE I. Plate parameters used for sample calculations.

Description of parameter	Symbol	Value
Mass density	ρ_p	2800 kg/m^3
Thickness	h_p	0.005 m
Elastic modulus	E_p	6.9×10^{10} Pa
Loss factor	η_p	0.001
Poisson's ratio	ν_p	0.3
Length	L	0.8 m

where k_f is the flexural wave number of a thin plate

$$k_f=\left(\frac{\rho_p\omega^2 h_p}{D}\right)^{1/4}. \tag{58}$$

The **F** matrix in Eq. (57) can be factored explicitly as

$$\mathbf{\Lambda}(\omega)=\begin{bmatrix}k_f&0&0&0\\0&ik_f&0&0\\0&0&-k_f&0\\0&0&0&-ik_f\end{bmatrix} \tag{59}$$

and

$$\mathbf{V}(\omega)=\begin{bmatrix}1&1&1&1\\k_f&ik_f&-k_f&-ik_f\\k_f^2&-k_f^2&k_f^2&-k_f^2\\k_f^3&-ik_f^3&-k_f^3&ik_f^3\end{bmatrix}. \tag{60}$$

Referring to Eq. (4) the acoustic loading is applied via

$$\boldsymbol{\gamma}=[0\quad 0\quad 0\quad 1/D]^T. \tag{61}$$

The boundary conditions in (54) and (55) are written in the form of Eq. (5) with

$$\mathbf{M}(\omega)=\begin{bmatrix}1&0&0&0\\0&0&1&0\\0&0&0&0\\0&0&0&0\end{bmatrix}, \tag{62}$$

$$\mathbf{N}(\omega)=\begin{bmatrix}0&0&0&0\\0&0&0&0\\1&0&0&0\\0&i\omega\mu&1&0\end{bmatrix}, \tag{63}$$

and $\hat{\boldsymbol{\epsilon}}(\omega)=\mathbf{0}$. In addition, Tables I and II show the properties and dimensions of the plate and the ambient fluid used in the numerical simulations.

TABLE II. Ambient fluid properties used for sample calculations.

Description of parameter	Symbol	Value
Mass density of fluid	ρ_0	1.18 kg/m^3
Speed of sound	c	330 m/s

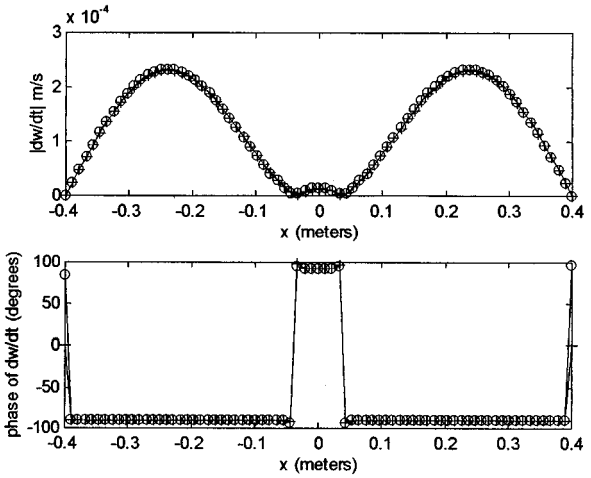


FIG. 3. Normal velocity of the plate as a function of x calculated at 200 Hz. The ○ markers are from the analytic solutions and the + markers are from the finite element model.

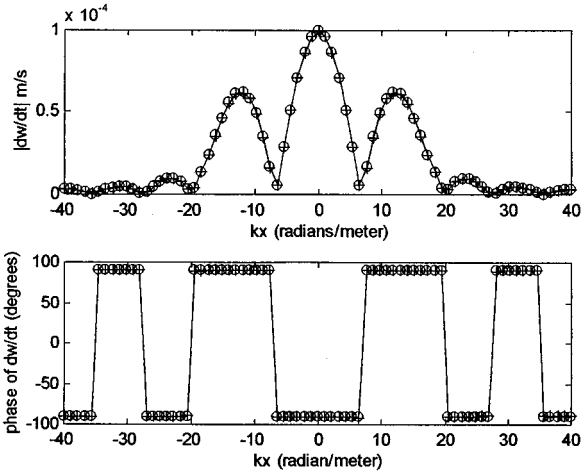


FIG. 4. The wave number transform $\tilde{w}(k_x,\omega)$ of the normal velocity calculated at 200 Hz.

A. Finite element verification

To demonstrate the accuracy of the analytic solutions, numerical calculations from both the analytic solutions and the NASTRAN finite element model are compared. The test case presented here assumes simply supported boundary conditions, $\theta_i=-\pi/2$, $p_i=1$, and $\omega=200$ Hz. Other test cases for the boundary actuation and the point drive can be found in Ref. 5.

The NASTRAN finite element model consists of 74 beam elements with properties consistent with the Mindlin theory flexural rigidity of the plate strip described in Table I. The mesh was chosen using the rule of thumb used for finite element method vibration analyses that one flexural wave length in a platelike structure be modeled with at least 12 elements for an accurate solution. Noting that the flexural wave length is $\lambda_f=2\pi/k_f$ and using (58) can determine an adequate mesh for the strip structure.

Figure 3 shows the normal velocity of the plate as a function of x calculated at 200 Hz. The ○ markers are from the analytic solution (44) and the + markers are from the finite element model. In the sequel, ○ and + will be used to distinguish the results from the analytic solutions and the finite element solutions. Figure 3 shows that the analytic solutions predict the normal velocity with good accuracy.

Figure 4 shows the wave number transform $\tilde{w}(k_x,\omega)$ of the normal velocity calculated at 200 Hz. The analytic predictions are calculated from (47) and the finite element predictions are calculated by substituting the normal velocity from the finite element model into the definition of the wave number transform [i.e., (13)] and integrating numerically. Looking at Fig. 4, one sees peaks in the magnitude plot at $k_x\approx\pm 11\ \mathrm{m}^{-1}$. These peaks roughly correspond to the infinite plate flexural wave number k_f defined in (58). Evaluation of (58) at 200 Hz yields $k_f\approx 12.9\ \mathrm{m}^{-1}$. The boundary conditions imposed on the strip shift the peaks in the wave number spectrum. Therefore, the peaks at $k_x\approx\pm 11\ \mathrm{m}^{-1}$ correspond to a standing wave contribution of wavelength ($\lambda_x=2\pi/k_x$) of about 0.57 m. This feature is consistent with the normal velocity distribution shown in Fig. 3, in which a strong sinusoidal component of $\lambda_x\approx 0.57$ m is readily apparent. In addition to these peaks, there is a strong component at $k_x=0$. Examining (47), one would predict that the first term of (47) would contribute a strong component at $k_x=-k\cos\theta_i$. In this case, $\theta_i=-\pi/2$ and $\cos\theta_i=0$ resulting in a peak at $k_x=0$. If this type of wave number transform is typical of that of the surrounding frequencies, one would predict that the strip subjected to this type of load will radiate more at the lower frequencies and the coincidence phenomena will be less pronounced, as can be seen by the radiated power given in (12).

Figure 5 shows the far-field radiated pressured $\hat{p}_t$ calculated at 200 Hz and a constant range of 30 m. The analytic predictions are calculated from (48). The finite element calculations use the Helmholtz–Kirchhoff integral equation for flat strip in polar coordinates[5]

$$\hat{p}(r_0,\theta_0)=\frac{\rho_0\omega}{2}\int_{-L/2}^{+L/2}\hat{w}(x)\cdot H_0^{(2)}(k\sqrt{r_0^2+x^2-2r_0x\cos\theta_0})\,dx, \qquad (64)$$

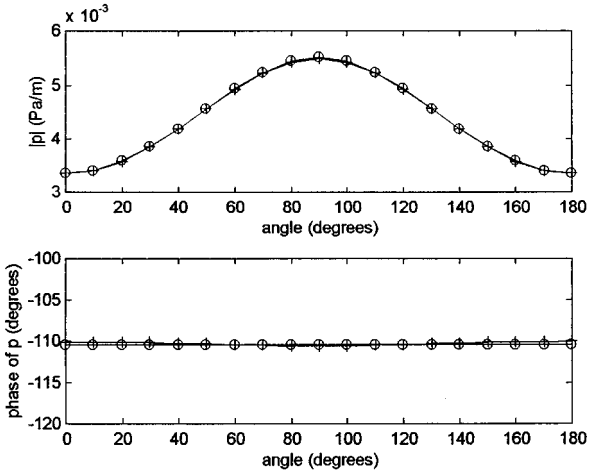


FIG. 5. The far-field radiated pressured $\hat{p}_t$ calculated at 200 Hz and a constant range of 30 m.

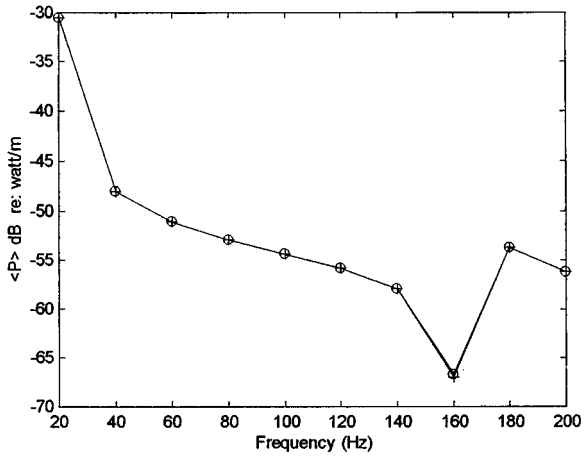


FIG. 6. The radiated power calculated from 20 to 200 Hz with a frequency increment of 20 Hz.

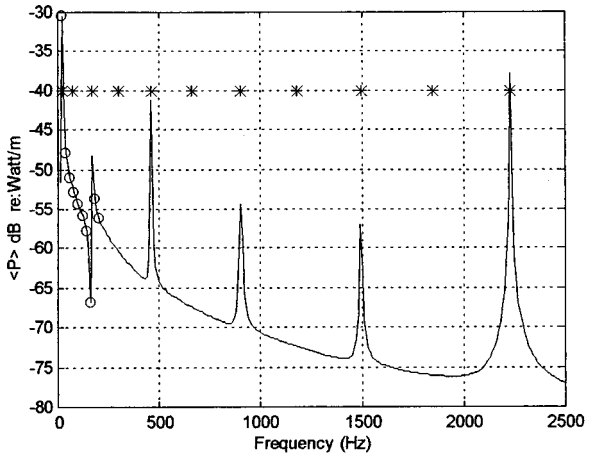


FIG. 7. The radiated power calculated from 10 to 2500 Hz with a frequency increment of 10 Hz.

where $H_0^{(2)}$ is the Hankel function of the second kind and the normal velocity $\hat{w}(x)$ is obtained from the finite element simulations. The agreement between the analytic solution and the finite element simulation is very good.

Figure 6 shows the radiated power from 20 to 200 Hz at 20-Hz frequency increments. The analytic predictions are calculated from (49). To facilitate finite element predictions, the radial component of the fluid velocity corresponding to the pressure in (64) is found as

$$\hat{v}_r(r_0,\theta_0)=\frac{-ik}{2}\int_{-L/2}^{+L/2}\hat{w}(x)\,\frac{(r_0-x\,\cos\,\theta_0)}{\sqrt{r_0^2+x^2-2r_0x\,\cos\,\theta_0}}\times H_1^{(2)}(k\sqrt{r_0^2+x^2-2r_0x\,\cos\,\theta_0})dx, \quad (65)$$

where $\hat{w}(x)$ is obtained from the finite element simulations. Then the radiated power is calculated from

$$\langle \mathbf{P}\rangle=\frac{1}{2}\,\mathrm{Re}\left\{\int_0^{\pi}\hat{p}(r_0,\theta)\cdot\hat{v}_r^{*}(r_0,\theta)r_0d\theta\right\}. \quad (66)$$

The limited frequency range and coarse frequency mesh in Fig. 6 are used, because the Helmholz–Kirchhoff formulation for the radiated power is computationally intensive. Using the analytic solution (49), a fairly fine frequency mesh and high-frequency range can be calculated with relative ease. Figure 7 shows the radiated power calculated from 10 to 2500 Hz with a frequency increment of 10 Hz, where the * indicates the natural frequencies of the simply supported strip.

The agreement between the analytic solutions and finite element simulations are in general good. Nevertheless, the following issues need to be discussed. First of all, the coincidence effect is not pronounced because of the strong radiation at about 20 Hz, which has been suggested by the wave number transform in Fig. 4. Moreover, only the odd numbered modes are excited, because the acoustic loading in (52) is symmetric. Finally, significant acoustic radiation occurs at about 20 and 2200 Hz as indicated in Fig. 7. If the acoustic loading in (52) was interpreted as the excitation of an incident wave, the average power input of the incident wave would be about -30 dB re: Ws/m, which is about the level of the transmitted power at about 20 and 2200 Hz. It is apparently contradicting the assumption that the radiated pressure is much smaller than the incident and reflected components of the pressure field around the structure. Therefore, the results at 20 and 2200 Hz should be interpreted as the harmonic acoustic loading $2p_i\,\exp(i\omega t+ikx\cos\theta_i)$ rather than an incident plane pressure wave.

B. Active control of acoustic radiation

To demonstrate the application of the analytic solution to active noise control, consider the acoustic radiation from the strip structure subjected to an acoustic loading in (52) with $\theta_i=-45°$. The asymptotically stable boundary control scheme is given in (55). Equation (49) was used to calculate the radiated power for the properties and dimensions described in Tables I and II. Note that the boundary condition (55) involves an arbitrary control gain μ and the optimal value of μ resulting in maximized noise reduction remains unknown at this stage. Therefore, it takes a trial and error process to determine the optimal value of μ.

The most efficient way to perform this sort of trial and error experimentation is to first calculate and store boundary independent values of K, $\mathbf{\Omega}$, Ξ, and κ in Eq. (49) at each frequency. Evaluating the effect of changing the boundary conditions, in this case by changing the value of μ, is then a relatively quick and simple computational procedure. It consists of reevaluating the constant $\boldsymbol{\phi}$ via Eq. (46) at each frequency and then performing a simple expansion with the other terms in (49) to obtain the total transmitted power result.

By varying the parameter μ as described above and then examining the transmitted sound power spectrum, there seemed to be a single optimum value $\mu\approx 33\ \mathrm{m}^{-1}\,\mathrm{s}$ for which the greatest reduction of sound power was achieved. Figures 8, 9, and 10 compare the this controlled case with the original passive system ($\mu=0$) at frequency intervals of 0–1 kHz, 1–2 kHz, and 2–3 kHz, respectively. According to (49), the radiated power that is totally independent of the boundary conditions is given by

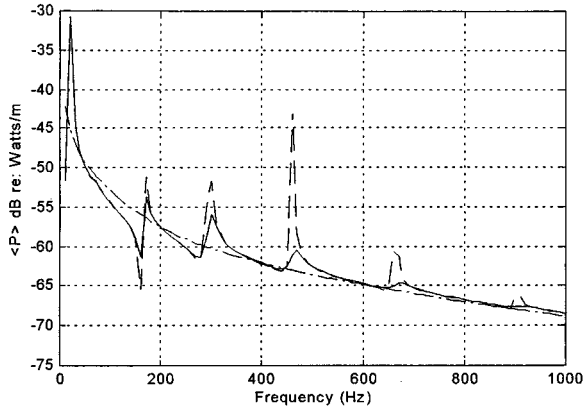


FIG. 8. The radiated power of the controlled case (solid line) compared with the original passive system (dashed line) at frequency intervals of 0–1 kHz; dash-dot line is the radiated power independent of boundary condition from (67).

$$\langle \mathbf{P}'(\omega,\theta_i)\rangle = \frac{\rho_0 c k}{4\pi} K(\omega,\theta_i). \tag{67}$$

For comparison, $\langle \mathbf{P}'(\omega,\theta_i)\rangle$ is also shown in Figs. 8–10 as the dash-dot lines.

Figures 8, 9, and 10 show that the transmitted power peaks of the passive case correspond to the natural frequencies of the simply supported structure given by

$$f_n = \frac{n^2\pi}{2L^2}\sqrt{\frac{D}{\rho_p h_p}}. \tag{68}$$

Looking at Figs. 8–10, one can see that there is a reduction at all of the resonance frequencies by this control scheme. The peaks of the controlled case are slightly shifted in frequency from those of the original simply supported strip. The levels first begin to decrease with increasing frequency and then increase until the coincidence frequency ($f_c \approx 2300$ Hz), after which the levels are decreasing. The boundary independent term, described in Eq. (67), appears to form a lower bound of the transmitted power, contributing most at the lower frequencies. Since the control scheme affects only the boundary dependent terms, it is relatively ineffective at the lower frequencies. Past the point where the boundary independent component stops contributing substantially, the peak closest to the coincidence frequency is the highest, as one would expect. Note that the greatest transmitted sound power reduction is obtained at 2230 Hz that corresponds to the strip natural frequency that is closest to the coincidence frequency.

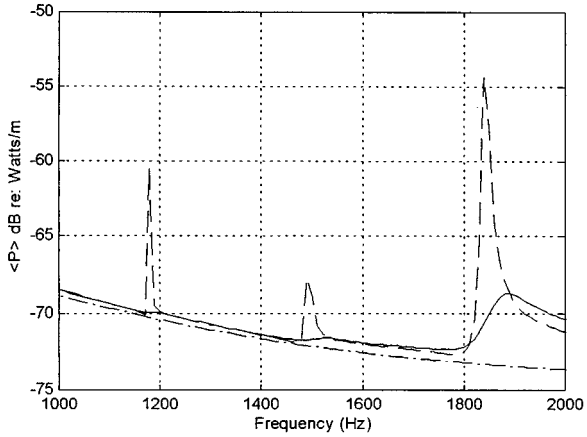


FIG. 9. The radiated power of the controlled case (solid line) compared with the original passive system (dashed line) at frequency intervals of 1–2 kHz; dash-dot line is the radiated power independent of boundary condition from (67).

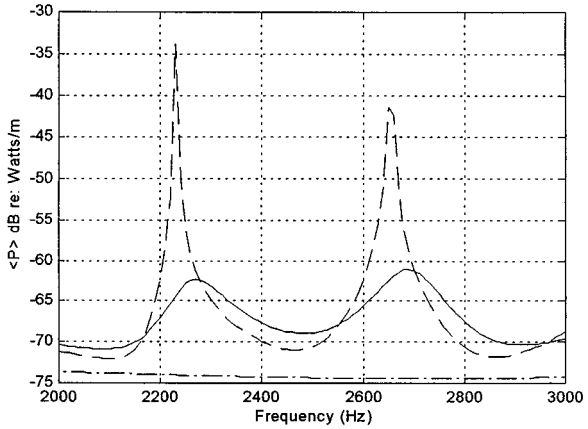


FIG. 10. The radiated power of the controlled case (solid line) compared with the original passive system (dashed line) at frequency intervals of 2–3 kHz; dash-dot line is the radiated power independent of boundary condition from (67).

It is also very easy to look at the far-field pressure radiation and it is interesting to see how the far-field pressure varies with this sort of control/boundary condition. Figure 11 shows the directivity pattern for the strip with and without control calculated via (48). Calculations are performed at a range of 30 m and a driving frequency of 2230 Hz with $\mu \approx 33\ \mathrm{m}^{-1}\,\mathrm{s}$. According to (48), the radiated pressure that is independent of the boundary conditions is

$$\hat{p}_t'(r,\theta,\omega) \approx \rho_0 \omega \kappa L \frac{e^{-i(kr-\pi/4)}}{\sqrt{2\pi kr}} \cdot \frac{\sin[kL(\cos\theta+\cos\theta_i)/2]}{kL(\cos\theta+\cos\theta_i)/2}, \quad 0<\theta<\pi,\ kr\gg 1,\ \text{and}\ r\gg L. \tag{69}$$

Figure 11 indicates a change in the directivity pattern from the uncontrolled to the controlled case, with the character of the controlled case reflecting some of the significant characteristics associated with the boundary independent component of the response. Most apparent is the peak in pressure directivity of the controlled case at 135°, corresponding to that for the boundary independent component of the response. Theoretically, an infinite plate with this loading will be transparent at 135°.

VI. CONCLUDING REMARKS

The solutions presented in this paper allow for a relatively flexible, quick, and easy way of performing numerical experiments to study the effect of changing various parameters and modeling assumptions of the strip on the radiated sound power. It is especially easy to examine the effect of changing the boundary conditions of the strip since the most

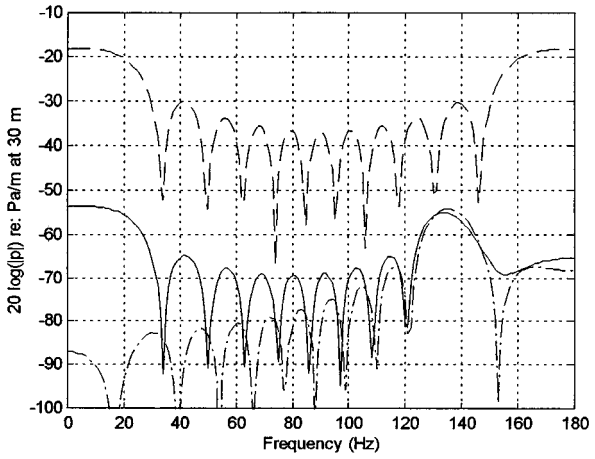


FIG. 11. Far-field pressure at range of 30 m and a driving frequency of 2230 Hz with $\mu \approx 33\ \mathrm{m}^{-1}\ \mathrm{s}$; original passive system (dash line), controlled system (solid line), and the boundary independent term from (69) (dash-dot line).

time consuming portion of computing the solution is independent of the boundary conditions. Hence, this portion of the solution need only be computed once and the total solution is then obtained by calculating the simple expressions dependent on the boundary conditions and then multiplying this expression by the other component of the solution to get the full solution.

It is very easy to incorporate active control at a boundary using this method as illustrated in the last section. Solutions for other specific excitations may also be acquired by using the method outlined in this paper, and taking advantage of the solution to the function I_{mn} that enabled the closed form expressions for the radiated power to be obtained.

The solutions presented in this paper are based on the fact that the normal displacement of the strip can be expressed as a linear combination of exponential functions of x with a finite number of terms.[4] Some of these exponential functions may have arguments that have large real parts. This is illustrated quite clearly when the eigenvalues of the **F** matrix corresponding to a thin plate were calculated in closed form in the last section. In (60), some of the eigenvalues of the **F** matrix are predominantly real numbers. These values will increase in magnitude with increasing frequency. Looking at the general expressions given in Eqs. (6) and (9), avoiding significant numerical round off requires that

$$\max(|\Re e(\lambda_m)|)\cdot L < 4.6\cdot n, \tag{70}$$

where λ_m is an eigenvalue of the **F** matrix and n is the floating point precision available. It was found that for some structures, such as a constrained layer system composed of elastic plates and a particular viscoelastic material, some of the resulting eigenvalues of the **F** matrix were so large that this condition was exceeded even at low frequencies.

APPENDIX A: EVALUATION OF $I_{mn}(\delta_0)$

This Appendix is to derive the closed form solution (23) of the function $I_{mn}(\delta_0)$ defined as

$$I_{mn}(\delta_0)=\int_{-k}^{+k}\frac{e^{ik_x\delta_0}}{(\lambda_m^*-ik_x)(\lambda_n+ik_x)\sqrt{k^2-k_x^2}}\,dk_x. \tag{A1}$$

The first step is to make the following change of variable

$$k_x=k\ \sin\ \theta,\quad \sqrt{k^2-k_x^2}=+k\ \cos\ \theta,\quad -\frac{\pi}{2}\leqslant\theta\leqslant\frac{\pi}{2}. \tag{A2}$$

Substitution of Eq. (A2) into (A1) gives

$$I_{mn}(\delta_0)=\int_{-\pi/2}^{+\pi/2}\frac{e^{ik\delta_0\ \sin\ \theta}}{(\lambda_m^*-ik\ \sin\ \theta)(\lambda_n+ik\ \sin\ \theta)}\,d\theta. \tag{A3}$$

Alternatively, the following change of variable

$$k_x=k\ \sin\ \theta,\quad \sqrt{k^2-k_x^2}=-k\ \cos\ \theta,\quad \frac{\pi}{2}\leqslant(-\theta)\leqslant\frac{3\pi}{2} \tag{A4}$$

gives

$$I_{mn}(\delta_0)=\int_{\pi/2}^{3\pi/2}\frac{e^{ik\delta_0\ \sin\ \theta}}{(\lambda_m^*-ik\ \sin\ \theta)(\lambda_n+ik\ \sin\ \theta)}\,d\theta. \tag{A5}$$

Therefore, the addition of Eqs. (A3) and (A5) implies that

$$I_{mn}(\delta_0)=\frac{1}{2}\int_0^{2\pi}\frac{e^{ik\delta_0\ \sin\ \theta}}{(\lambda_m^*-ik\ \sin\ \theta)(\lambda_n+ik\ \sin\ \theta)}\,d\theta. \tag{A6}$$

The limits of integration in Eq. (A6) are a total cycle. This form will allow the integral to eventually be cast into a contour on the complex plane that will be very easy to evaluate.

The term $e^{ik\delta_0\ \sin\ \theta}$ in Eq. (A6) can be expanded as a Neumann series[9]

$$e^{ik\delta_0\ \sin\ \theta}=\sum_{\nu=1}^{+\infty}J_{-\nu}(k\delta_0)\cdot e^{-i\nu\theta}+\sum_{\nu=0}^{+\infty}J_\nu(k\delta_0)\cdot e^{i\nu\theta}. \tag{A7}$$

Substitution of Eq. (A7) into Eq. (A6) gives

$$I_{mn}(\delta_0)=I_{mn}^{(1)}(\delta_0)+I_{mn}^{(2)}(\delta_0), \tag{A8}$$

where

$$I_{mn}^{(1)}(\delta_0)=\frac{1}{2}\int_0^{2\pi}\frac{\Sigma_{\nu=1}^{+\infty}J_{-\nu}(k\delta_0)\cdot e^{-i\nu\theta}}{(\lambda_m^*-ik\ \sin\ \theta)(\lambda_n+ik\ \sin\ \theta)}\,d\theta \tag{A9}$$

and

$$I_{mn}^{(2)}(\delta_0)=\frac{1}{2}\int_0^{2\pi}\frac{\Sigma_{\nu=0}^{+\infty}J_{\nu}(k\delta_0)\cdot e^{i\nu\theta}}{(\lambda_m^*-ik\ \sin\ \theta)(\lambda_n+ik\ \sin\ \theta)}\,d\theta. \tag{A10}$$

To find a closed form expression for Eq. (A9), define the following change of variable along the unit circle

$$z=-e^{-i\theta} \tag{A11}$$

and apply

$$J_{-\nu}(k\delta_0)=(-1)^\nu J_\nu(k\delta_0) \tag{A12}$$

to (A9), resulting in

$$I_{mn}^{(1)}(\delta_0) = \frac{2i}{k^2} \oint_{|z|=1} \frac{z \cdot \Sigma_{\nu=1}^{\infty} J_\nu(k\delta_0) \cdot z^\nu}{(z-P_m^{(1)})(z-P_m^{(2)})(z-P_n^{(1)})(z-P_n^{(2)})}\, dz, \tag{A13}$$

where the contour integration is along the unit circle in a counterclockwise sense and $P_m^{(1)}$, $P_m^{(2)}$, $P_n^{(1)}$, and $P_n^{(2)}$ are poles defined in (24) and (25). Similarly, the change of variable

$$z = e^{i\theta} \tag{A14}$$

will transform (A10) into

$$I_{mn}^{(2)}(\delta_0) = \frac{2i}{k^2} \oint_{|z|=1} \frac{z \cdot \Sigma_{\nu=0}^{\infty} J_\nu(k\delta_0) \cdot z^\nu}{(z-P_m^{(1)})(z-P_m^{(2)})(z-P_n^{(1)})(z-P_n^{(2)})}\, dz. \tag{A15}$$

Substitution of (A13) and (A15) into (A8) gives

$$I_{mn}(\delta_0) = \frac{2i}{k^2} \oint_{|z|=1} \frac{z \cdot [J_0(k\delta_0) + 2 \cdot \Sigma_{\nu=1}^{\infty} J_\nu(k\delta_0) \cdot z^\nu]}{(z-P_m^{(1)})(z-P_m^{(2)})(z-P_n^{(1)})(z-P_n^{(2)})}\, dz. \tag{A16}$$

This form of $I_{mn}(\delta_0)$, a contour integral, is evaluated as $2\pi i$ times the sum of the residues of Eq. (A16) given the poles ($P_m^{(1,2)}$, and $P_n^{(1,2)}$) that lie within the region $|z|=1$ on the complex plane.[10] A sufficient condition to prevent the poles from being on the unit circle is not to have purely imaginary λ_m and λ_n.

Finally, it is also easy to verify that

$$I_{mn}(0) = \frac{2i}{k^2} \oint_{|z|=1} \frac{z\, dz}{(z-P_m^{(1)})(z-P_m^{(2)})(z-P_n^{(1)})(z-P_n^{(2)})} \tag{A17}$$

and

$$I_{mn}(-\delta_0) = I_{nm}^{*}(\delta_0). \tag{A18}$$

APPENDIX B: ASYMPTOTIC STABILITY OF THE STRIP STRUCTURE

This Appendix is to apply Lyapuvonv's direct method[11] to prove that the strip structure is asymptotically stable under the control law given in (55). Consider the following Lyapunov functional;

$$F(x,t) = \int_{-L/2}^{+L/2} \left[D\left(\frac{\partial^2 w}{\partial x^2}\right)^2 + \rho_s h_s \left(\frac{\partial w}{\partial t}\right)^2 \right] dx. \tag{B1}$$

The time derivative of (B1) can be reduced solely in terms of the boundary conditions[8]

$$\frac{dF}{dt} = \frac{\partial^2 w}{\partial x^2}\frac{\partial^2 w}{\partial x\, \partial t}\bigg|_{x=-L/2}^{x=+L/2} - \frac{\partial^3 w}{\partial x^3}\frac{\partial w}{\partial t}\bigg|_{x=-L/2}^{x=+L/2}. \tag{B2}$$

Substitution of the boundary conditions (54) and (55) to (B2) results in

$$\frac{dF}{dt} = -\mu \left\{ \frac{\partial^2 w}{\partial x\, \partial t}\bigg|_{x=-L/2}^{x=+L/2} \right\}^2 < 0 \tag{B3}$$

for all time t and $\mu > 0$. The asymptotic stability is guaranteed in the Lyapunov sense because $dF/dt < 0$.

[1] L. Cremer and M. Heckl, *Structure-Borne Sound: Structural Vibrations and Sound Radiation at Audio Frequencies* (Springer-Verlag, New York, 1973).

[2] C. Guigou and C. R. Fuller, ''Active control of sound radiation from a semi-infinite elastic beam with a clamped edge,'' J. Sound Vib. **168**, 507–523 (1993).

[3] W. T. Baumann, W. R. Saunders, and H. H. Robertshaw, ''Active suppression of acoustic radiation from impulsively excited structures,'' J. Acoust. Soc. Am. **90**, 3202–3208 (1991).

[4] B. Yang and C. A. Tan, ''Transfer functions of one-dimensional distributed parameter systems,'' ASME J. Appl. Mech. **59**, 1009–1014 (1992).

[5] C. Ting, ''Solutions for the Vibro-Acoustic Response of Strip Structures,'' Master's thesis, University of Washington, Seattle, WA, 1995.

[6] I. Y. Shen, ''Stability and controllability of Euler–Bernoulli beams with intelligent constrained layer,'' ASME J. Vib. Acoust. **118**, 70–77 (1996).

[7] M. Junger and D. Feit, *Sound, Structures, and Their Interaction* (MIT, Cambridge, MA, 1986).

[8] T. Baily and J. Hubbard, ''Distributed piezoelectric-polymer active vibration control of a cantilever beam,'' J. Guidance **8**, 605–611 (1985).

[9] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1972).

[10] C. Wyle and L. Barret, *Advanced Engineering Mathematics* (McGraw-Hill, New York, 1982).

[11] C. Dym, *Mechanics of Elastic Stability* (Noordohoff International, Leyden, 1974).

Modal impulse response and wave vector–time domain methods to evaluate transient radiation from axisymmetric planar radiators

Jie Sun[a] and Peter R. Stepanishen
Department of Ocean Engineering, University of Rhode Island, Narragansett, Rhode Island 02882

(Received 27 January 1997; accepted for publication 9 April 1997)

The modal impulse response method and the wave vector–time domain method are two of the existing methods to evaluate the transient acoustic field from an axisymmetric planar radiator. Both methods are investigated here and the wave vector–time domain method is used to derive the modal impulse response for the normal velocity distribution of an arbitrary axisymmetric source. Representations of the transient acoustic field in the space–time domain and the wave vector–time domain are linked through the Fourier–Hankel transform. A piston is first addressed as a result of its simplicity which serves as a benchmark case; a Bessel shaded aperture is then addressed as a result of its importance. Clearly, the piston case is a special case of the spatial Bessel shading cases. Numerical results are presented for both cases to illustrate the procedures in using the modal impulse response and the wave vector–time domain methods. © *1997 Acoustical Society of America.* [S0001-4966(97)01308-8]

PACS numbers: 43.40.Rj [CBB]

INTRODUCTION

Enormous effort has been put into evaluating the acoustic transient field from an axisymmetric planar radiator. Starting from a spatially uniform piston-type radiator,[1–5] it is now more interesting to study a spatially varying radiator or aperture which is typical of a multielement annular array. If the space and time variables of a source are separable, the impulse response approach can be used to investigate the space–time pressure field.

For a general space–time nonseparable planar source, two of the existing methods to project the acoustic field are the modal impulse response method[1,6] and the wave vector–time domain method (also called the angular spectrum method).[7–10] The relationship between the two methods has been investigated by Stepanishen *et al.*[11] In particular, the modal impulse response method provides a bridge, which makes it possible to represent a space–time nonseparable source in terms of space–time separable sources, through the use of a modal series expansion,[6] so that a modal impulse response can be defined for each modal term in the series expansion.

In the present paper, the wave vector–time domain method will be used to explicitly derive the modal impulse response which has already been formulated by Stepanishen.[6] The wave vector–time domain representation of the modal impulse response will also be derived. Following Stepanishen,[1,6] the results for the cases of a piston and a Bessel shaded aperture will be presented. The space–time domain and the wave vector–time domain are linked through the Fourier–Hankel transforms. The relationship between the two methods, including the relationship between the operations in the different domains, will be explicitly established.

[a]Currently at Department of Geological Sciences, Lincoln Field Building, Box 1846, Brown University, Providence, RI 02912.

I. THEORY

Consider an axisymmetric planar source that is in contact with a fluid as shown in Fig. 1. Normalized space, time coordinates (ρ,φ,Z,τ) are used in the following development where the characteristic length is the radius of the source radiator and the characteristic time is the acoustic travel time across the radius. The basic problem of interest is to evaluate the acoustic transient pressure field in the upper half-space $Z\geqslant 0$ due to a specified normal velocity at the source.

In the cylindrical coordinate shown in the Fig. 1, the source is centered at the origin and in the $Z=0$ plane. The position of a source element is denoted as $(\rho_s,\varphi_s,0)$, while the position of an acoustic field point is denoted as (ρ,φ,Z), the relative azimuth angle between the source element and the acoustic field point is $\theta=\varphi_s-\varphi$. As a result of the axial symmetry of the source, the acoustic field is also axisymmetric, so φ can be set to be zero, thus $\theta=\varphi_s$.

A. Space–time domain

The acoustic wave equation for the velocity potential $\phi(\rho,Z,\tau)$ can be expressed as

$$\nabla^2\phi(\rho,Z,\tau)=\frac{\partial^2\phi(\rho,Z,\tau)}{\partial^2\tau}, \tag{1}$$

and the acoustic pressure $p(\rho,Z,\tau)$ can be expressed as

$$p(\rho,Z,\tau)=\frac{\partial\phi(\rho,Z,\tau)}{\partial\tau}. \tag{2}$$

The linearized momentum equation can then be used to relate the normal velocity $u(\rho,0,\tau)$ of the planar source to the velocity potential via the following Neumann boundary condition

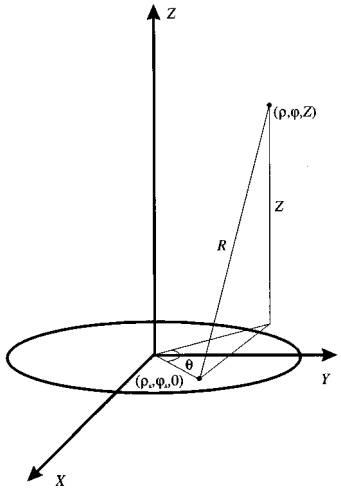


FIG. 1. The geometry of interest.

$$u(\rho,0,\tau)=-\left.\frac{\partial\phi(\rho,Z,\tau)}{\partial Z}\right|_{Z=\mathbf{0}}. \tag{3}$$

The solution to the boundary value problem of interest is well-known and can be expressed in a number of different forms. Perhaps the simplest form for the acoustic pressure is the following Rayleigh surface integral:

$$p(\rho,Z,\tau)=\int_{S_0}\frac{1}{2\pi R}\left.\frac{\partial u(\rho_s,0,\tau_s)}{\partial\tau_s}\right|_{\tau_s=\tau-R} dS_{\mathbf{0}}, \tag{4}$$

which follows from the Green's function solution of the boundary value problem where S_0 denotes the source area, and $R=\sqrt{\rho^2+\rho_s^2-2\rho\rho_s\cos\theta+Z^2}$.[6,12,13] It is implied in Eq. (4) that a spatial convolution and a temporal convolution have to be evaluated in order to determine the acoustic pressure field from the normal velocity distribution of the source.

B. Transform domains

For normalized time τ and its associated normalized angular frequency k_τ, the Fourier transform pair is defined as

$$F(k_\tau)=\int_{-\infty}^{\infty}f(\tau)e^{-jk_\tau\tau}\,d\tau,$$
$$f(\tau)=\frac{1}{2\pi}\int_{-\infty}^{\infty}F(k_\tau)e^{jk_\tau\tau}\,dk_\tau. \tag{5}$$

For normalized spatial coordinate X and Y and their associated normalized angular wave number k_X and k_Y, the Fourier transform pair is defined as

$$F(k_X,k_Y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f(X,Y)e^{-j(k_XX+k_YY)}\,dX\,dY,$$
$$f(X,Y)=\frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}F(k_X,k_Y)e^{j(k_XX+k_YY)}\,dk_X\,dk_Y. \tag{6}$$

Assuming axial symmetry, and defining $\rho^2=X^2+Y^2$ and $k_\rho^2=k_X^2+k_Y^2$, the Hankel transform pair can be derived directly following Eq. (6) as

$$F(k_\rho)=2\pi\int_{\mathbf{0}}^{\infty}f(\rho)J_{\mathbf{0}}(k_\rho\rho)\rho\,d\tau,$$
$$f(\tau)=\frac{1}{2\pi}\int_{\mathbf{0}}^{\infty}F(k_\rho)J_{\mathbf{0}}(k_\rho\rho)k_\rho\,dk_\rho. \tag{7}$$

So the transform pair connecting the space–time domain and the wave number–frequency domain is

$$\hat{F}(k_\rho,Z,k_\tau)=2\pi\int_{-\infty}^{\infty}\int_{\mathbf{0}}^{\infty}\rho f(\rho,Z,\tau)J_0(\rho k_\rho)\times e^{-jk_\tau\tau}\,d\rho\,d\tau,$$
$$f(\rho,Z,\tau)=\frac{1}{(2\pi)^3}\int_{-\infty}^{\infty}\int_{0}^{\infty}k_\rho\hat{F}(k_\rho,Z,k_\tau)\times J_{\mathbf{0}}(\rho k_\rho)e^{jk_\tau\tau}\,dk_\rho\,dk_\tau. \tag{8}$$

The Fourier–Hankel transform of Eqs. (1)–(3) leads to the following expressions in the wave vector–frequency domain

$$\frac{d^2\hat{\Phi}(k_\rho,Z,k_\tau)}{dZ^2}+(k_\tau^2-k_\rho^2)\hat{\Phi}(k_\rho,Z,k_\tau)=0, \tag{9}$$

$$\hat{P}(k_\rho,Z,k_\tau)=jk_\tau\hat{\Phi}(k_\rho,Z,k_\tau), \tag{10}$$

$$\hat{U}(k_\rho,k_\tau)=-\left.\frac{\partial\hat{\Phi}(k_\rho,Z,k_\tau)}{\partial Z}\right|_{Z=\mathbf{0}}. \tag{11}$$

After solving Eq. (9) subject to the radiation condition and the boundary condition in Eq. (11), the transform of the velocity potential can be expressed as

$$\hat{\Phi}(k_\rho,Z,k_\tau)=\frac{\hat{U}(k_\rho,k_\tau)}{j\sqrt{k_\tau^2-k_\rho^2}}e^{-j\sqrt{k_\tau^2-k_\rho^2}Z}, \tag{12}$$

and the transform of the acoustic pressure can be expressed as

$$\hat{P}(k_\rho,Z,k_\tau)=\hat{U}(k_\rho,k_\tau)\frac{k_\tau e^{-j\sqrt{k_\tau^2-k_\rho^2}Z}}{\sqrt{k_\tau^2-k_\rho^2}}. \tag{13}$$

A projection operator can thus be defined in the wave vector–frequency domain

$$\hat{G}(k_\rho,Z,k_\tau)=\frac{e^{-j\sqrt{k_\tau^2-k_\rho^2}Z}}{j\sqrt{k_\tau^2-k_\rho^2}}. \tag{14}$$

The convolution property of the Fourier transform can now be employed to express the acoustic pressure in the wave vector–time domain, by taking the inverse Fourier transform of Eq. (13) with respect to k_τ, i.e.,

$$P(k_\rho,Z,\tau)=\begin{cases}\dfrac{\partial U(k_\rho,\tau)}{\partial\tau}\underset{\tau}{\otimes}G(k_\rho,Z,\tau)\\ U(k_\rho,\tau)\underset{\tau}{\otimes}\dfrac{\partial G(k_\rho,Z,\tau)}{\partial\tau},\end{cases} \tag{15}$$

where $\underset{\tau}{\otimes}$ denotes a convolution with respect to τ, the projection operator in the wave vector–time domain is

$$G(k_\rho,Z,\tau)=J_0(k_\rho\sqrt{\tau^2-Z^2}), \tag{16}$$

and its time derivative is

$$\frac{\partial G(k_\rho,Z,\tau)}{\partial \tau}=\frac{\partial J_0(k_\rho\sqrt{\tau^2-Z^2})}{\partial \tau}$$

$$=\delta(\tau-Z)-k_\rho^2\tau\,\frac{J_1(k_\rho\sqrt{\tau^2-Z^2})}{k_\rho\sqrt{\tau^2-Z^2}}. \tag{17}$$

The acoustic pressure $p(\rho,Z,\tau)$ can then be determined from an inverse Hankel transform of $P(k_\rho,Z,\tau)$ with respect to k_ρ.

II. MODAL IMPULSE RESPONSE

Following Stepanishen,[6] consider a modal series expansion of an arbitrary space–time normal velocity distribution, of which each term can be regarded as a space–time separable normal velocity distribution, i.e.,

$$u(\rho,0,\tau)=\sum_n u_n(\rho,0,\tau)=\sum_n \psi_n(\rho)u_n(\tau), \tag{18}$$

where $\psi_n(\rho)$ represents a complete set of orthonormal functions and $u_n(\tau)$ is the corresponding velocity coefficient for each mode. Consider now only one term in the series expansion which is a space–time separable normal velocity distribution, i.e.,

$$u(\rho,0,\tau)=\psi(\rho)u(\tau) \tag{19}$$

whose corresponding Hankel transform is

$$U(k_\rho,0,\tau)=\Psi(k_\rho)u(\tau), \tag{20}$$

and the subscript n is dropped for simplicity. After substituting Eq. (20) into Eq. (15) and regrouping, the Hankel transform of the impulse response can be defined as

$$H(k_\rho,Z,\tau)=\Psi(k_\rho)G(k_\rho,Z,\tau), \tag{21}$$

so that Eq. (15) becomes

$$P(k_\rho,Z,\tau)=\begin{cases}\dfrac{\partial u(\tau)}{\partial \tau}\underset{\tau}{\otimes}H(k_\rho,Z,\tau)\\ u(\tau)\underset{\tau}{\otimes}\dfrac{\partial}{\partial \tau}H(k_\rho,Z,\tau).\end{cases} \tag{22}$$

By taking the inverse Hankel transform of both $P(k_\rho,Z,\tau)$ and $H(k_\rho,Z,\tau)$, and using the convolution property of the Hankel transform,[14] an impulse response $h(\rho,Z,\tau)$ can be defined as such that

$$p(\rho,Z,\tau)=\frac{\partial}{\partial \tau}\,[u(\tau)\underset{\tau}{\otimes}h(\rho,Z,\tau)],$$

where

$$h(\rho,Z,\tau)=\psi(\rho)\underset{\rho,\theta}{\otimes}g(\rho,Z,\tau)$$

$$=\begin{cases}\displaystyle\int_0^\infty\int_0^{2\pi}\psi(R_s)g(\rho_s,Z,\tau)\rho_s\,d\rho_s\,d\theta\\ \displaystyle\int_0^\infty\int_0^{2\pi}\psi(\rho_s)g(R_s,Z,\tau)\rho_s\,d\rho_s\,d\theta,\end{cases} \tag{23}$$

where $R_s=\sqrt{\rho^2+\rho_s^2-2\rho\rho_s\cos\theta}$, and $\underset{\rho,\theta}{\otimes}$ denotes a two-dimensional spatial convolution with respect to ρ and θ in the cylindrical coordinate.

It is well-known from a Hankel transform table[14] that

$$g(\rho,Z,\tau)=\frac{1}{2\pi}\,\frac{\delta(\rho-\sqrt{\tau^2-Z^2})}{\sqrt{\tau^2-Z^2}}, \tag{24}$$

which is the projection operator in the space–time domain. After substituting Eq. (24) into the upper expression of Eq. (23) and using the shifting property of the delta function, the impulse response can then be written in the form

$$h(\rho,Z,\tau)=\frac{1}{2\pi}\int_0^{2\pi}\psi(R_0)\,d\theta, \tag{25}$$

where

$$R_0=(\tau^2-Z^2+\rho^2-2\sqrt{\tau^2-Z^2}\rho\,\cos\,\theta)^{1/2}. \tag{26}$$

Consider now a spatial shading function of the normal velocity source distribution

$$\psi(\rho),\quad 0\leqslant\rho\leqslant 1, \tag{27}$$

i.e., $\psi(\rho)$ is zero for $\rho>1$. It is thus apparent that in Eq. (25), R_0 has to satisfy

$$0\leqslant R_0\leqslant 1 \tag{28}$$

and it then follows from Eqs. (26) and (28) that

$$0\leqslant(\sqrt{\tau^2-Z^2}-\rho)^2\leqslant\tau^2-Z^2+\rho^2-2\sqrt{\tau^2-Z^2}\rho\,\cos\,\theta$$

$$\leqslant(\sqrt{\tau^2-Z^2}+\rho)^2\leqslant 1. \tag{29}$$

Upon examining Eq. (29), it is obvious that

$$\tau\geqslant Z \tag{30}$$

and the $\rho\geqslant 1$ and $\rho\leqslant 1$ cases have to be handled separately.

For $\rho\geqslant 1$, $(\sqrt{\tau^2-Z^2}+\rho)^2$ is always greater than or equal to 1. So $(\sqrt{\tau^2-Z^2}-\rho)^2\leqslant 1$ is first considered, which leads to

$$\sqrt{(Z^2+(1-\rho)^2}\leqslant\tau\leqslant\sqrt{(Z^2+(1+\rho)^2}. \tag{31}$$

It is equally obvious from $\tau^2-Z^2+\rho^2-2\sqrt{\tau^2-Z^2}\rho\cos\theta\leqslant 1$ that the following inequality holds

$$-\arccos\left(\frac{\tau^2-Z^2+\rho^2-1}{2\sqrt{\tau^2-Z^2}\rho}\right)$$

$$\leqslant\theta\leqslant\arccos\left(\frac{\tau^2-Z^2+\rho^2-1}{2\sqrt{\tau^2-Z^2}\rho}\right). \tag{32}$$

For $\rho\leqslant 1$, if $(\sqrt{\tau^2-Z^2}+\rho)^2\leqslant 1$, i.e.,

$$\tau\leqslant\sqrt{(Z^2+(1-\rho)^2}, \tag{33}$$

there is no limitation on θ, so

$$-\pi\leqslant\theta\leqslant\pi, \tag{34}$$

and if $(\sqrt{\tau^2-Z^2}+\rho)^2\geqslant 1$, i.e.,

$$\tau\geqslant\sqrt{(Z^2+(1-\rho)^2}, \tag{35}$$

then $(\sqrt{\tau^2-Z^2}-\rho)^2\leqslant 1$ and $\tau^2-Z^2+\rho^2-2\sqrt{\tau^2-Z^2}\rho\cos\theta\leqslant 1$, which also yields Eq. (31) and (32).

Combining Eqs. (31)–(35), these limits are exactly the same as presented by Stepanishen.[6] Thus Eq. (25) leads to the same expression for the impulse response as previously derived:[6]

$$h(\rho,Z,\tau)=\frac{1}{\pi}\int_{\mathbf{0}}^{\theta_e(R)}\psi(\rho_s(R,\theta))\Big|_{R=\tau}\,d\theta,$$

$$R_{\min}<\tau\leqslant R_{\max} \tag{36}$$

where $R=\sqrt{\rho^2+\rho_s^2-2\rho\rho_s\cos\theta+Z^2}$, and
for $\rho\geqslant 1$,

$$\rho_s(R,\theta)=[\rho^2+R^2-Z^2-2\rho(R^2-Z^2)^{1/2}\cos\theta]^{1/2},$$

$$\theta_e(R)=\arccos\left(\frac{R^2-Z^2+\rho^2-1}{2\rho\sqrt{R^2-Z^2}}\right),\quad R_{\min}<\tau\leqslant R_{\max},$$

$$R_{\min}=\sqrt{Z^2+(\rho-1)^2},\quad R_{\max}=\sqrt{Z^2+(\rho+1)^2}, \tag{37}$$

for $\rho<1$,

$$\rho_s(R,\theta)=[\rho^2+R^2-Z^2-2\rho(R^2-Z^2)^{1/2}\cos\theta]^{1/2},$$

$$\theta_e(R)=\begin{cases}\pi, & Z<\tau\leqslant R_{\min}\\ \arccos\left(\dfrac{R^2-Z^2+\rho^2-1}{2\rho\sqrt{R^2-Z^2}}\right), & R_{\min}<\tau\leqslant R_{\max},\end{cases}$$

$$R_{\min}=\sqrt{Z^2+(1-\rho)^2},\quad R_{\max}=\sqrt{Z^2+(1+\rho)^2}. \tag{38}$$

III. SPECIAL CASES

A. A piston

Consider a piston where the spatial shading function can be expressed as

$$\psi(\rho)=1,\quad 0\leqslant\rho\leqslant 1. \tag{39}$$

This case can be regarded as a special case of the spatial Bessel shading cases where $ka_s=0$. The impulse responses for the piston[1] can easily be obtained, which are recited here as a benchmark for completeness.

For $\rho>1$,

$$h(\rho,Z,\tau)=\begin{cases}0, & \tau\leqslant R_{\min},\\ \dfrac{1}{\pi}\arccos\left(\dfrac{\tau^2-Z^2+\rho^2-1}{2\sqrt{\tau^2-Z^2}\rho}\right), & R_{\min}<\tau\leqslant R_{\max},\\ 0, & R_{\max}\leqslant\tau\end{cases} \tag{40}$$

and for $\rho\leqslant 1$,

$$h(\rho,Z,\tau)=\begin{cases}0, & \tau\leqslant Z,\\ 1, & Z<\tau\leqslant R_{\min},\\ \dfrac{1}{\pi}\arccos\left(\dfrac{\tau^2-Z^2+\rho^2-1}{2\sqrt{\tau^2-Z^2}\rho}\right), & R_{\min}<\tau\leqslant R_{\max},\\ 0, & R_{\max}\leqslant\tau.\end{cases} \tag{41}$$

The Hankel transform of $\psi(\rho)$ can simply be obtained

$$\Psi(k_\rho)=\frac{2\pi J_1(k\rho)}{k_\rho}=\pi(J_0(k_\rho)+J_2(k_\rho)). \tag{42}$$

Although it is very difficult to derive the Hankel transform for the impulse response from Eqs. (40) and (41), its Hankel transform can easily be obtained in an analytical form using the wave vector–time domain method, by combining Eqs. (16) and (42) to obtain

$$H(k_\rho,Z,\tau)=2\pi J_0(k_\rho\sqrt{\tau^2-Z^2})\,\frac{J_1(k_\rho)}{k_\rho}$$
$$=\pi J_0(k_\rho\sqrt{\tau^2-Z^2})(J_0(k_\rho)+J_2(k_\rho)). \tag{43}$$

B. A Bessel shaded aperture

Consider a Bessel shaded aperture where the spatial shading function $\psi(\rho)$ can be expressed as

$$\psi(\rho)=J_0(ka_s\rho),\quad 0\leqslant\rho\leqslant 1. \tag{44}$$

It is noted that ultrasonic transducers with such a spatial shading have received considerable interest as a result of their near-field properties.

From Stepanishen,[6] the impulse response for this special case can be written as
for $\rho\geqslant 1$

$$h(\rho,Z,\tau)=\frac{\psi_e(\tau)}{\pi}\sum_{m=0}^{\infty}\epsilon_m J_m(ka_s\rho)J_m[ka_s(\tau^2-Z^2)^{1/2}]$$
$$\times\left(\frac{\sin m\psi_e(\tau)}{m\psi_e(\tau)}\right),\quad R_{\min}<\tau<R_{\max}, \tag{45}$$

and for $\rho<1$

$$h(\rho,Z,\tau)=\begin{cases}J_0(ka_s\rho)J_0[ka_s(\tau^2-Z^2)^{1/2}], & Z<\tau\leqslant R_{\max},\\ \dfrac{\psi_e(\tau)}{\pi}\displaystyle\sum_{m=0}^{\infty}\epsilon_m J_m(ka_s\rho)J_m[ka_s(\tau^2-Z^2)^{1/2}]\\ \quad\times\left(\dfrac{\sin m\psi_e(\tau)}{m\psi_e(\tau)}\right), & R_{\min}<\tau\leqslant R_{\max},\end{cases} \tag{46}$$

where

$$\psi_e(\tau)=\begin{cases}\pi\\ \arccos\left(\dfrac{\tau^2-Z^2+\rho^2-1}{2\sqrt{\tau^2-Z^2}\rho}\right).\end{cases} \tag{47}$$

The Hankel transform of $\psi(\rho)$ can simply be obtained[15]

$$\Psi(k_\rho)=\begin{cases}\dfrac{2\pi}{k_\rho^2-ka_s^2}(k_\rho J_0(ka_s)J_1(k_\rho)-ka_s J_0(k_\rho)J_1(ka_s)),,\\ \pi((J_1(ka_s))^2+(J_0(ka_s))^2), \quad ka_s=k_\rho.\\ \quad ka_s\neq k_p.\end{cases} \tag{48}$$

Although it is very difficult to derive the Hankel transform for the impulse response from Eqs. (45) and (46), its Hankel transform can also easily be obtained in an analytical form using the wave vector–time domain method, by combining Eqs. (16) and (48) to obtain

$$H(k_\rho, Z, \tau)$$

$$= J_0(k_\rho \sqrt{\tau^2 - Z^2}) \times \begin{cases} \dfrac{2\pi}{k_\rho^2 - ka_s^2}\,(k_\rho J_0(ka_s)J_1(k_\rho) - ka_s J_0(k_\rho)J_1(ka_s)), \\ \pi((J_1(ka_s))^2 + (J_0(ka_s))^2), \quad ka_s = k_\rho. \\ ka_s \neq k_p, \end{cases} \tag{49}$$

IV. NUMERICAL RESULTS

Numerical results for a piston and a Bessel shaded aperture are now presented. The spatial shading with the structural wave numbers $ka_s = 0.0$ and $ka_s = 10.17$ are shown in Fig. 2, while their on-axis peak pressures for the harmonic case are shown in Fig. 3. For $ka_s = 0.0$, the typical piston-type near- to far-field features are seen. For $ka_s = 10.17$, a typical near-field oscillation of the on-axis peak pressure is first seen and then a much steeper decay of the on-axis peak pressure is observed. The steeper decay is of course the result of the zero volumetric velocity which produces a far-field null associated with the fact that the pressure decays as Z^{-n} where $n > 1$. It is also noted that the near- to far-field transition distance of the piston is a factor of 10 greater than that of the Bessel shaded aperture.

Consider now a pulsed excitation where the velocity function used in the present study is a four cycle sine wave:

$$u(\tau) = \begin{cases} 0, & \tau \leqslant 0, \\ \sin(ka\tau), & 0 < \tau \leqslant 8\pi/ka \\ 0, & \tau > 8\pi/ka. \end{cases} \tag{50}$$

The normal velocity distribution for both cases are shown in Figs. 4 and 5, respectively.

The acoustic field at many different parallel planes has been numerically evaluated from the specified normal velocity distribution at $Z = 0$, using both the modal impulse response and the wave vector–time domain methods. However, as examples, only the numerical results at two parallel planes, i.e., $Z = 0.0$ and 4.0 are presented. For modal impulse response method, the upper expression of Eq. (22) was used, namely, the temporal derivative was performed on the time-dependent velocity function rather than on the impulse response. Similarly, for the wave vector–time domain method, the upper expression of Eq. (15) was used, namely, the temporal derivative was performed on the time-dependent velocity function rather than on the projection operator.

Consider first the piston case where $ka_s = 0.0$. Figures 6 and 7 show two impulse responses at two different Z's, with Figs. 8 and 9 showing their corresponding acoustic pressure field evaluated through the modal impulse response method. Figures 10 and 11 show the same acoustic pressure field but evaluated through the wave vector–time domain method.

The same computations are then repeated for the Bessel shading case where $ka_s = 10.17$. Figures 12 and 13 show two impulse responses at two different Z's, with Figs. 14 and 15 showing their corresponding acoustic pressure field evaluated through the modal impulse response method. Figures 16 and 17 show the same acoustic pressure field but evaluated through the wave vector–time domain method.

By examining the impulse responses, it is observed that for a fixed Z, the impulse response for the piston where $ka_s = 0.0$ can be regarded as the envelope of the impulse response for $ka_s > 0.0$, i.e., an impulse response for $ka_s > 0.0$ is a modulation of the envelope with a Bessel function. This fact has been partially observed by Satyapal[16,17] in terms of the temporal pulse width of the impulse response. In order to compute the field for $Z \gg 1$ and $\rho \ll 1$, it is noted that the temporal pulse width of the impulse response becomes very short, i.e., $1/2Z$. Very fine temporal resolution is thus needed, particularly when ka_s gets large. In order to minimize numerical errors, undersampling of the Bessel function has to be avoided.

By examining the acoustic field evaluated through the two different methods, it is clear that they agree with each other reasonably well. In the piston case, the spatial sharp edges, which are visible in the acoustic pressure field obtained through the modal impulse response method, have been rounded off in the acoustic field obtained through the wave vector–time domain method, due to the truncation of the wave-number spectrum.

In general, the wave vector–time domain method is very demanding in terms of resolution in both time and wave number. Fine temporal resolution is needed in order to minimize the numerical errors caused by the temporal convolution, while fine resolution is needed in wave number, in order to avoid undersampling of the Bessel function. At a projection plane that is far from the source and/or with a high wave number, the argument of the Bessel function will be

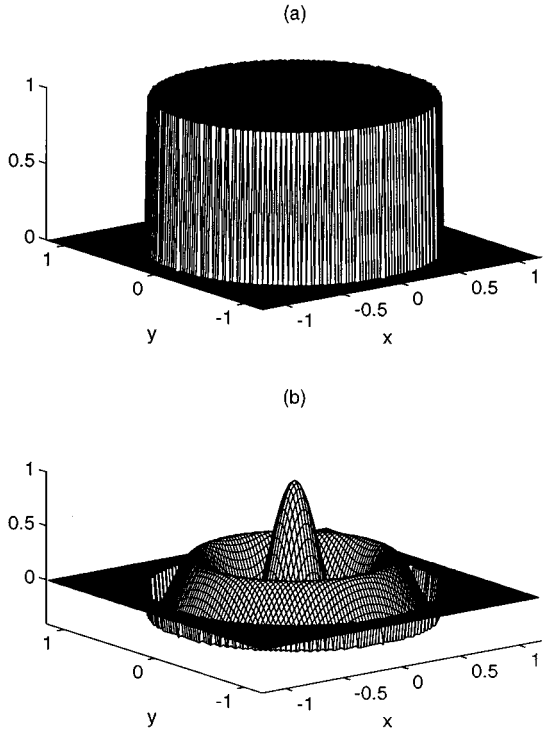


FIG. 2. Two spatial Bessel shading cases: (a) a piston: $ka_s = 0.0$, (b) a Bessel shaded aperture: $ka_s = 10.17$.

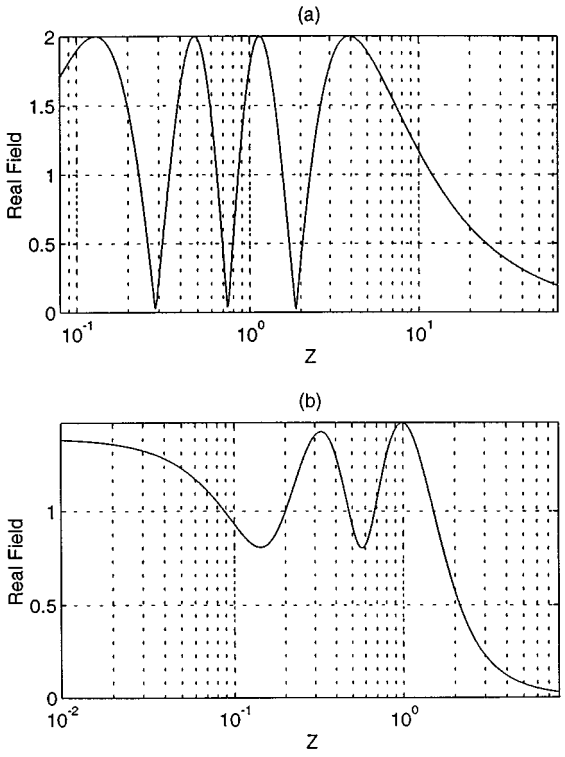


FIG. 3. The on-axis peak acoustic pressure square versus propagation distance Z for two spatial Bessel shading cases: (a) a piston: $ka_s=0.0$, (b) a Bessel shaded aperture: $ka_s=10.17$.

large, which causes a very rapid oscillation in the Bessel function. A trade-off thus has to be made between the reasonable length of a data array and enough samples for the Bessel function.

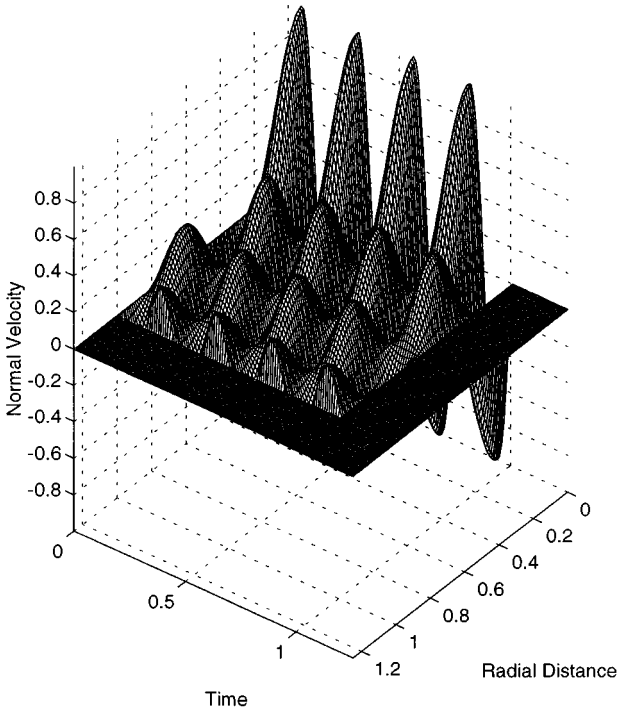


FIG. 5. The normal velocity distribution in space–time, when $ka_s=10.17$.

V. CONCLUSIONS

The modal impulse response and the wave vector–time domain methods have been presented to evaluate the acoustic transient pressure field from an axisymmetric planar radiator with a normal velocity distribution. The wave vector–time domain method has been investigated in such a way to shed more light on the definition of the modal impulse re-

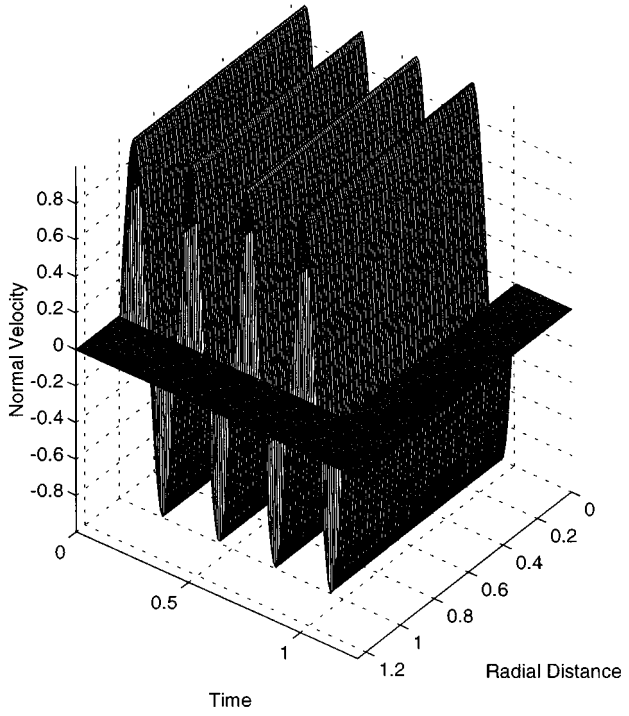


FIG. 4. The normal velocity distribution in space–time, when $ka_s=0.0$.

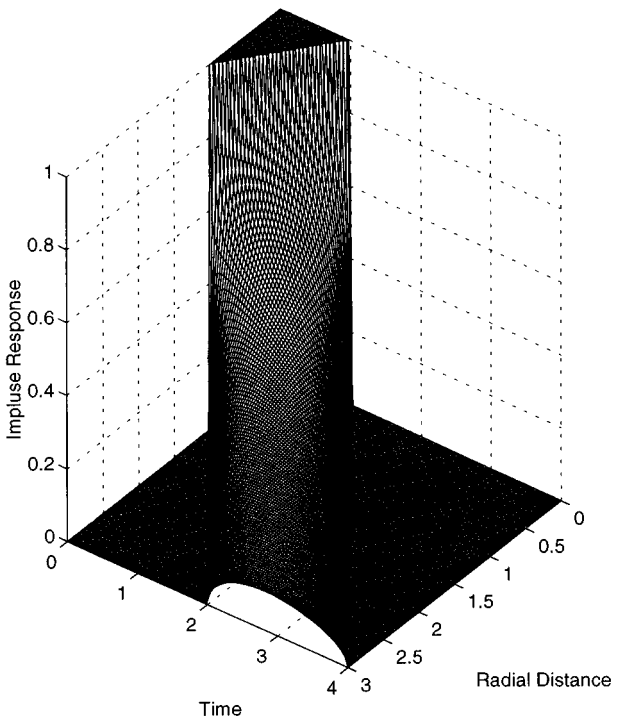


FIG. 6. The impulse response at $Z=0.0$ for a piston where $ka_s=0.0$.

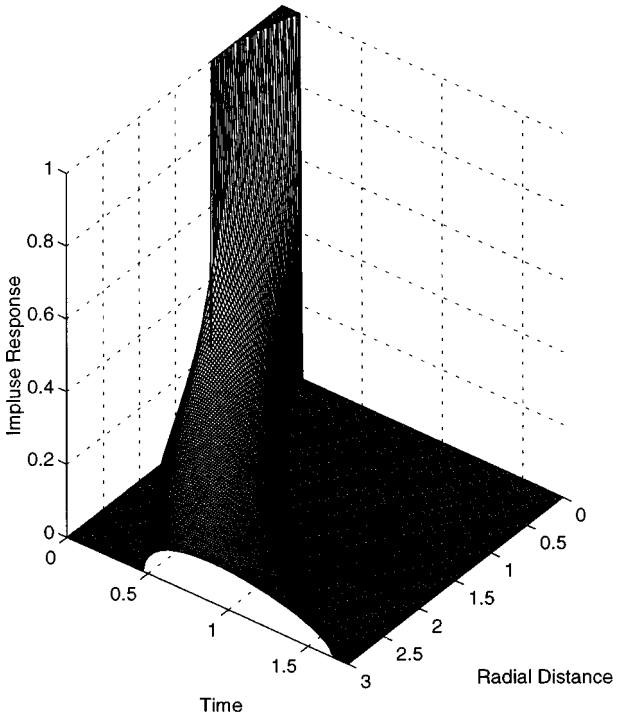


FIG. 7. The impulse response at $Z=4.0$ for a piston where $ka_s=0.0$.

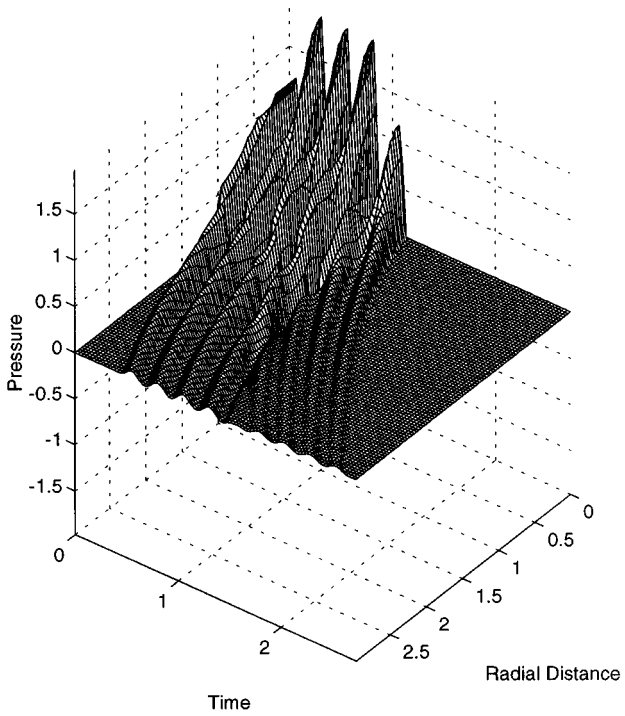


FIG. 9. Acoustic pressure field at $Z=4.0$, evaluated by using the modal impulse response method for a piston ($ka_s=0.0$).

sponse, i.e., the wave vector–time domain method has been presented as an alternative approach to derive the modal impulse response for an axisymmetric planar vibrator. The convolution property of the Hankel transform is used to connect the operations in the wave vector–time domain and the space–time domain, which provide more physical insight, e.g., (1) the commutative property for the two-dimensional convolution is equivalent to the coordinate transformation presented by Stepanishen,[6] and (2) for a finite aperture, one can analytically determine the upper and the lower bounds of

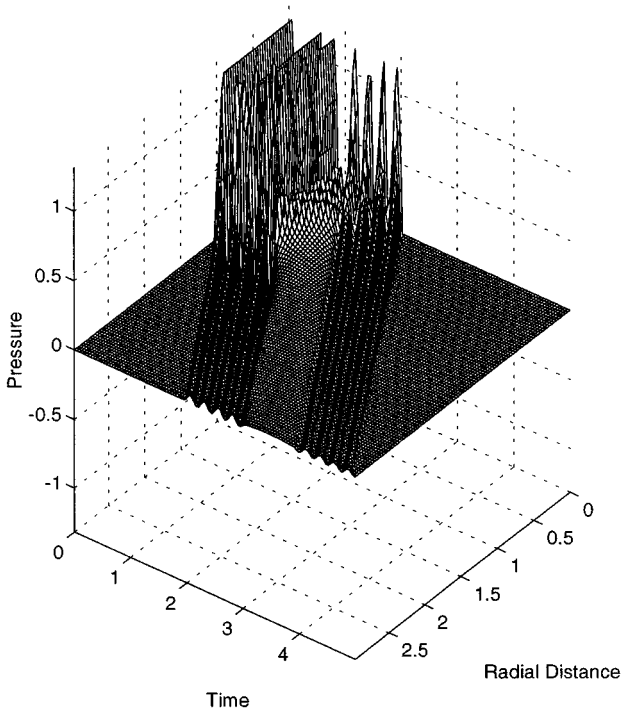


FIG. 8. Acoustic pressure field at $Z=0.0$, evaluated by using the modal impulse response method for a piston ($ka_s=0.0$).

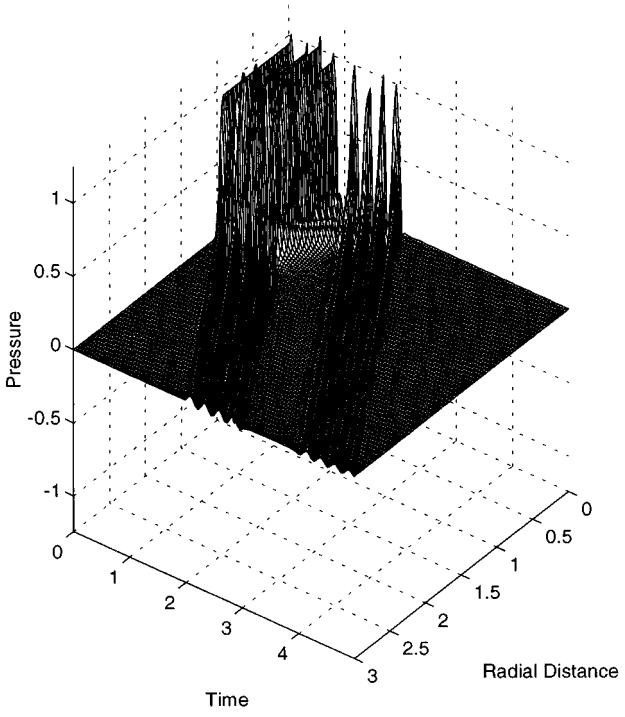


FIG. 10. Acoustic pressure field at $Z=0.0$, evaluated by using the wave vector–time domain method for a piston ($ka_s=0.0$).

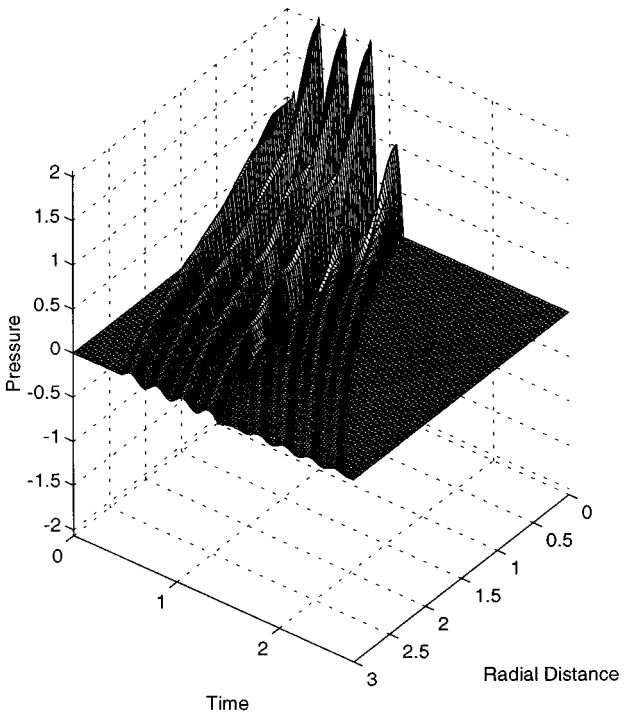


FIG. 11. Acoustic pressure field at $Z=4.0$, evaluated by using the wave vector–time domain method for a piston ($ka_s=0.0$).

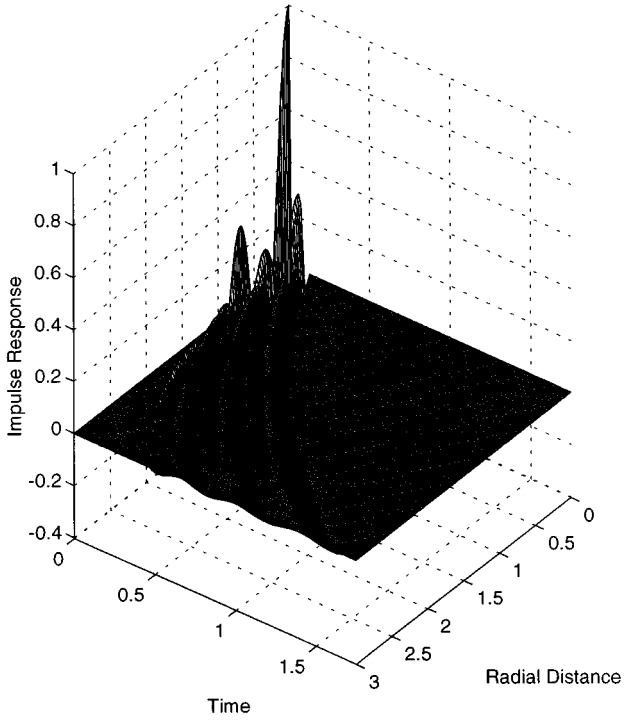


FIG. 13. The impulse response at $Z=4.0$ for a Bessel shaded aperture $ka_s=10.17$.

both time τ and azimuth angle θ by using the wave vector–time domain method, without reference to the actual geometry.

In order to illustrate the underlying relationship between the wave vector–time domain and the modal impulse response methods, the impulse response of a piston is first derived using the wave vector–time domain method, which is a special case of the Bessel beams, and also serves as the benchmark case. Then the impulse response of a circular radiator that has a spatial Bessel shading is derived using the wave vector–time domain approach. The Hankel transforms for both impulse responses have also been presented.

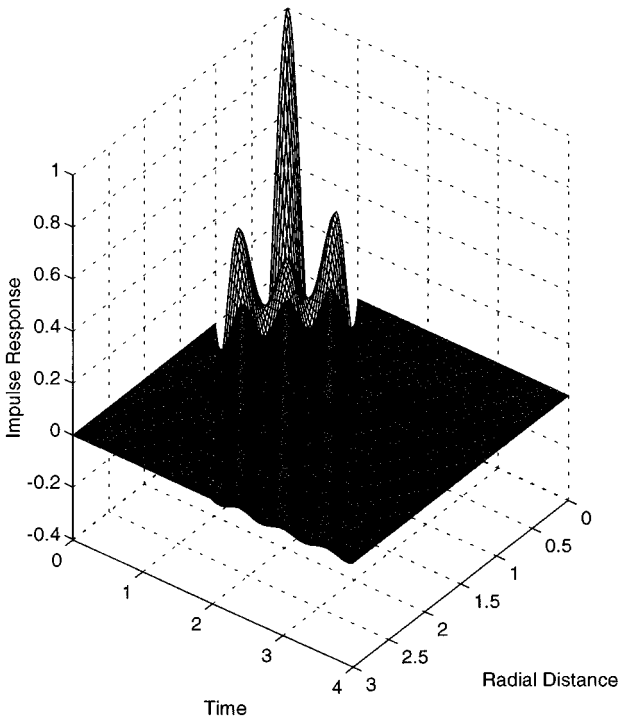


FIG. 12. The impulse response at $Z=0.0$ for a Bessel shaded aperture where $ka_s=10.17$.

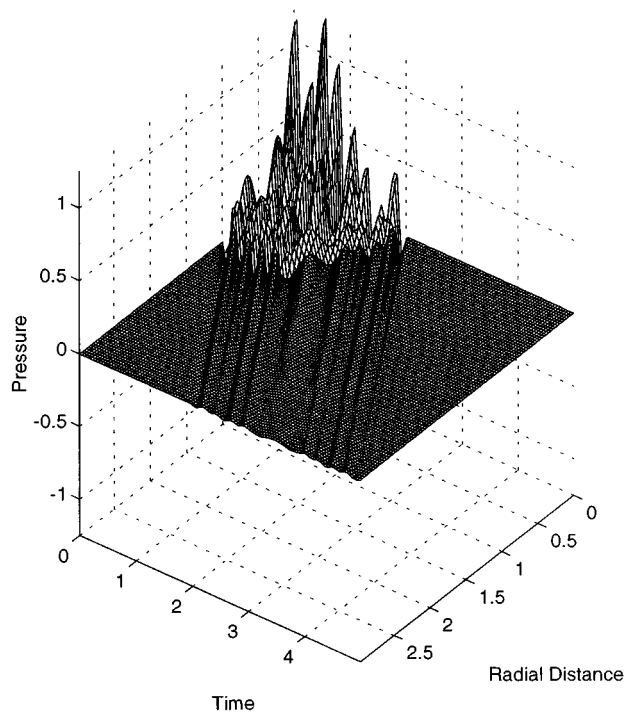


FIG. 14. Acoustic pressure field at $Z=0.0$, evaluated by using the modal impulse response method for a Bessel shaded aperture ($ka_s=10.17$).

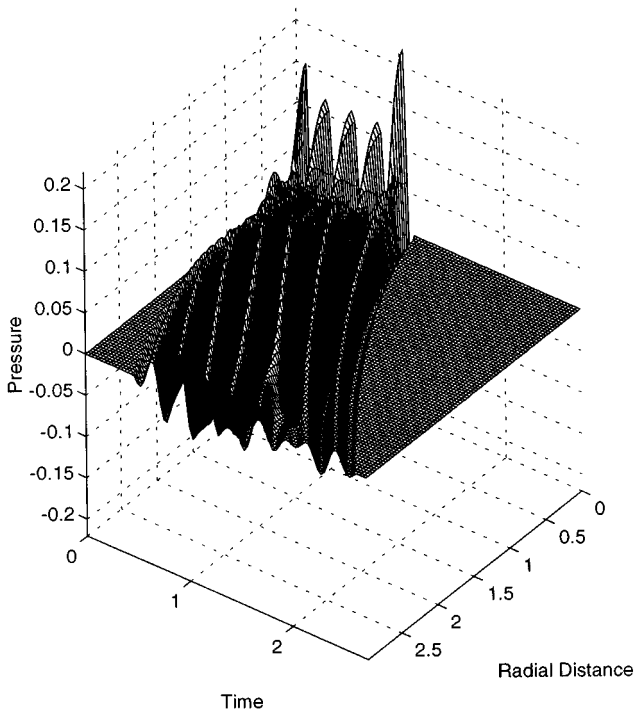


FIG. 15. Acoustic pressure field at $Z=4.0$, evaluated by using the modal impulse response method for a Bessel shaded aperture ($ka_s=10.17$).

Numerical results for a piston and a Bessel beam are presented to show that the acoustic field evaluated through the two different methods agree with each other reasonably well. A numerical quasifast Hankel transform has been developed. A close look at the numerical process suggests that the size of a resolution cell has to be very small in order to achieve certain acceptable numerical errors.

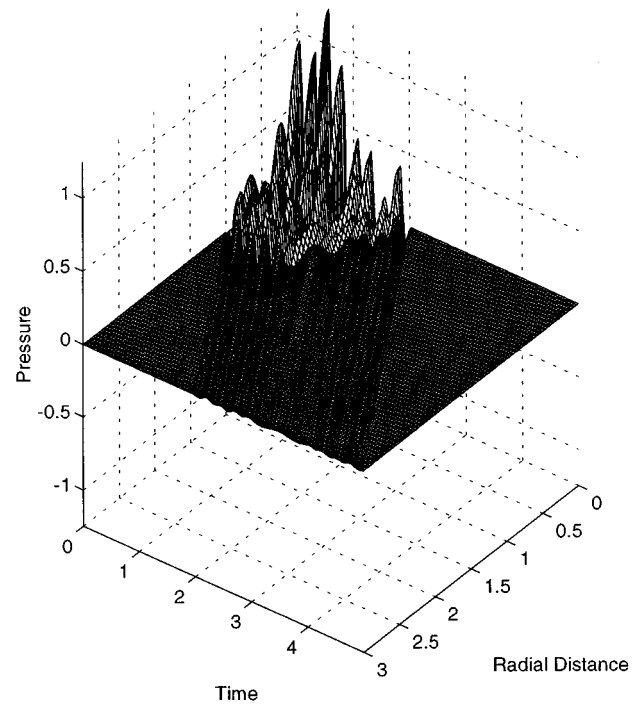


FIG. 16. Acoustic pressure field at $Z=0.0$, evaluated by using the wave vector–time domain method for a Bessel shaded aperture ($ka_s=10.17$).

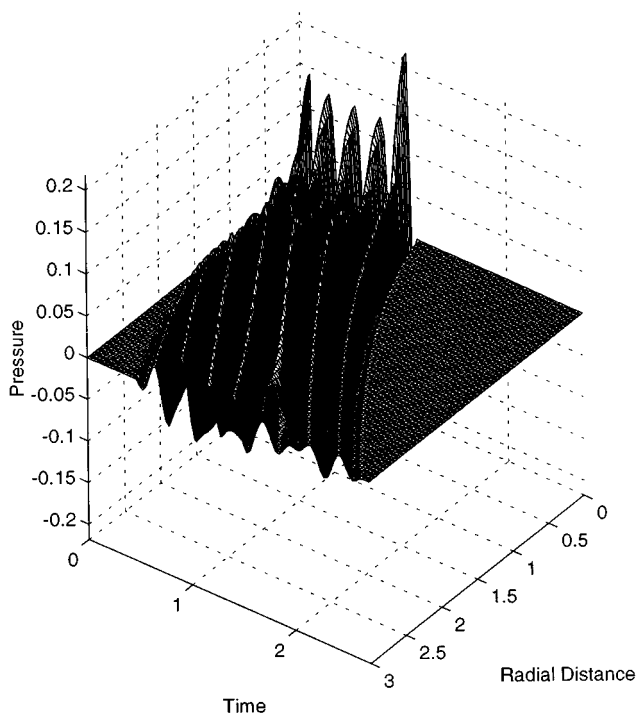


FIG. 17. Acoustic pressure field at $Z=4.0$, evaluated by using the wave vector–time domain method for a Bessel shaded aperture ($ka_s=10.17$).

The wave vector–time domain method is capable of projecting a space–time nonseparable source, while an impulse response can only be defined for a space–time separable source. The modal impulse response method provides a bridge, which makes it possible to represent a space–time nonseparable source in terms of space–time separable sources. Through the use of a modal series expansion, an impulse response can be defined for each modal term in the series expansion.

[1] P. R. Stepanishen, ''Transient radiation from pistons in an infinite planar baffle,'' J. Acoust. Soc. Am. **49**, 1629–1638 (1971).
[2] P. R. Stepanishen, ''Acoustic transients in the far field of a baffled circular piston using the impulse response approach,'' J. Sound Vib. **32**, 295–310 (1974).
[3] M. Greenspan, ''Piston radiator: Some extensions of the theory,'' J. Acoust. Soc. Am. **65**, 608–621 (1979).
[4] G. R. Harris, ''Transient field of a baffled planar piston having an arbitrary vibration amplitude distribution,'' J. Acoust. Soc. Am. **70**, 186–204 (1981).
[5] G. R. Harris, ''Review of transient field theory of a baffled planar piston,'' J. Acoust. Soc. Am. **70**, 10–20 (1981).
[6] P. R. Stepanishen, ''Acoustic transients from planar axisymmetric vibrators using the impulse response approach,'' J. Acoust. Soc. Am. **70**, 1176–1181 (1981).
[7] D. Guyomar and J. Powers, ''Boundary effects on transient radiation fields from vibrating surfaces,'' J. Acoust. Soc. Am. **77**, 907–915 (1985).
[8] D. Guyomar and J. Powers, ''Transient radiation from axially symmetric sources,'' J. Acoust. Soc. Am. **79**, 273–277 (1986).
[9] D. Guyomar and J. Powers, ''A fourier approach to diffraction of pulsed ultrasonic waves in lossless media,'' J. Acoust. Soc. Am. **82**, 354–359 (1987).
[10] M. Forbes, S. Letcher, and P. Stepanishen, ''A wave vector, time-domain method of forward projecting time-dependent pressure fields,'' J. Acoust. Soc. Am. **90**, 2782–2792 (1991).
[11] P. R. Stepanishen, M. Forbes, and S. Letcher, ''The relationship between the impulse response and angular spectrum methods to evaluate acoustic

transient fields,'' J. Acoust. Soc. Am. **90**, 2794–2798 (1991).
[12] J. W. Strutt (Lord Rayleigh), *The Theory of Sound* (Dover, New York, 1945).
[13] M. C. Junger and D. Feit, *Sound, Structures, and Their Interaction* (Acoustical Society of America, Woodbury, 1993).
[14] R. N. Bracewell, *The Fourier Transform and its Applications* (McGraw-Hill, New York, 1986), 2nd ed.
[15] L. Cremer, M. Heckl, and E. E. Ungar, *Structure-Borne Sound* (Springer-Verlag, New York, 1988), 2nd ed.
[16] S. Satyapal, ''Acoustic Transient Fields from Pulsed Ultrasonic Transducers and Arrays,'' Master's thesis, Department of Ocean Engineering, University of Rhode Island, 1992.
[17] S. Satyapal and P. R. Stepanishen, ''Space/time characteristics of localized transient fields from planar axisymmetric sources,'' J. Acoust. Soc. Am. **93**, 821–835 (1992).

Robust feedback control of flow-induced structural radiation of sound

Craig M. Heatwole, Matthew A. Franchek,[a] and Robert J. Bernhard
Ray W. Herrick Laboratories, School of Mechanical Engineering, Purdue University, West Lafayette, Indiana 47907-1077

(Received 15 March 1996; accepted for publication 4 March 1997)

A significant amount of the interior noise of aircraft and automobiles is a result of turbulent boundary layer excitation of the vehicular structure. In this investigation, the feasibility of active robust feedback control of the noise due to this type of excitation is studied. The structural sound radiation system investigated is comprised of a simply supported panel excited by turbulent flow. An analytical model of this system is developed to evaluate the robust feedback control system design strategy. The model consists of a modal description of the excitation of the plate by a turbulent boundary layer, a modal representation of the plate vibrations, and an acoustic model of the sound radiated from the plate. Parametric uncertainties of 1% variation in the natural frequencies and 5% variation in the damping ratio are considered. A frequency domain controller design approach is utilized to develop a robust controller. Performance objectives are specified which include output performance of the sound radiated by the plate and control effort constraints on the force applied to the plate by the controller. Plate acceleration feedback is used in an effort to reduce the sound-pressure level. Approximately 16 and 8 dB of sound-pressure level reduction were achieved at the first and third resonances, respectively. [S0001-4966(97)03607-2]

PACS numbers: 43.40.Vn [PJR]

INTRODUCTION

The structural radiation of sound resulting from a turbulent boundary layer is a major consideration in a variety of engineering applications. This phenomenon has been investigated by sonar, aircraft, and automobile designers. The turbulent flow pressure fluctuations from the propulsion of ships excite the sonar dome and induce a structural radiation of sound which interferes with the sonar receiver. As a result, the background noise is increased which limits sonar performance. Sound radiated into an aircraft interior is generated by skin panel vibration which is excited by such sources as the unsteady boundary layer flow over the fuselage, jet and propeller noise, compressor whine, and thrust reversal noise. The skin acts as a transducer and radiates sound into the interior. With the reduction of engine, turbine, and propeller noise and an increase of the speed of commercial aircraft, turbulent boundary layer noise has become a significant contributor to the interior sound field. Similarly, automotive engineers are increasingly concerned with the environment for passengers. Major advances have reduced the sound transmitted to the interior from the engine, transmission, and tires. As a result, the reduction of aeroacoustic noise has become a priority. This paper reports the results of an investigation of robust feedback control of turbulent flow-induced sound radiation. The case investigated is comprised of a simply supported rectangular flexible panel with turbulent flow excitation on one side.

Recently many adaptive feedforward control schemes have been utilized to actively control noise problems. However, these methods require either a deterministic excitation or a reference transducer which provides causal and coherent disturbance information. For turbulence driven panels, the turbulent boundary layer excitation is broadband and random in nature and therefore neither a deterministic source nor a highly coherent reference is available. Thus adaptive feedforward control is not a viable solution for this problem.

Relatively few studies have utilized robust feedback control for acoustic problems. For the case where the acoustic delay is small, Yang *et al.*[1,2] have used an H_∞ approach with sound-pressure level feedback to reduce sound inside a small reverberent enclosure. Robust control of structure-borne noise has been considered by Fluder and Kashani.[3] In their work, H_∞ and μ synthesis controllers were designed to reduce the amount of sound radiated from a plate. The excitation was from a point force and was broadband in nature. A model of the radiation of sound from the plate was not used for the controller design because the additional states required to model the sound radiation would substantially increase the controller order. Fluder and Kashani reported difficulty in obtaining a controller which satisfied both robust performance and robust stability. This is a consequence of the fact that the H_∞ and μ synthesis controller methodologies do not utilize beneficial plant dynamics. As a result, large control effort is required to achieve the performance objectives.

A feedback control study of the turbulence induced structural radiation of sound was presented by Thomas and Nelson.[4] In their study they developed a useful model for the turbulent flow excitation and utilized optimal control (linear quadratic regulator) as the feedback control technology. However, the authors point out that the use of this controller

[a]Corresponding author.

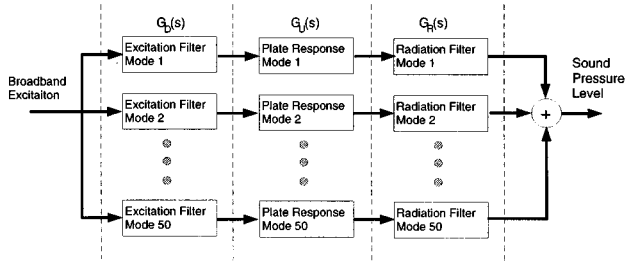


FIG. 1. Block diagram of an analytical model.

is not practical since it requires full state information. It is unlikely that the state information associated with the turbulent boundary layer or the sound radiation can be obtained through measurement or an accurate model. Furthermore, an optimal controller does not allow for uncertainty such as unmodeled modes and plant variations.

In this investigation a frequency domain robust controller design approach is used. The control approach is applied to a model of the turbulence induced structural sound radiation problem similar to that developed by Thomas and Nelson. Using this approach, the closed loop system is designed to be stable and achieve prespecified sound-pressure level reductions without saturating the control actuator for given system uncertainty.

I. ANALYTICAL MODEL

The analytical model for this investigation is used to predict the sound pressure resulting from the structural response of a simply supported panel due to a turbulent boundary layer excitation. The model is composed of three sets of modal transfer functions as shown in Fig. 1. The first set $[G_D(s)]$ is the generalized forces which capture the modal excitation of the plate due to the turbulent boundary layer. For a disturbance due to a point force excitation, these transfer functions are constants. However, for the turbulent boundary layer excitation they are shaping filters that give the appropriate modal force for an assumed white noise input. The second transfer function set $[G_U(s)]$ is used to represent the modal response of the plate. The third transfer function set $[G_R(s)]$ is used to model the relationship between the modal amplitudes of the plate and the sound-pressure level at a specified location. The complete analytical model uses 50 modes.

A. Turbulent flow model

Currently accurate models of turbulent flow fields do not exist. Despite years of research and a large body of literature devoted to the fluid dynamics of turbulent flow, this phenomenon remains an area of fundamental research. The most commonly used models of a turbulent flow field are statistical.

Statistical models of the turbulent wall pressure levels have been utilized since the early 1960s and are based on empirically obtained data. Pioneering work by Corcus led to statistical modeling of the wall pressure field. By assuming the flow field to be stationary and homogeneous, Corcos suggested that the wall pressure cross correlation is a function only of the spatial and temporal separation, not absolute position. This hypothesis assumes the turbulent boundary layer thickness is uniform and the mean pressure gradient is small.[5] With this assumption, Corcos postulated that the cross-spectral density of the wall pressure can be represented as

$$S_{ff}(\omega,\xi,\eta)=S_{pp}(w)A(\omega\xi/U_c)B(\omega\eta/U_c)\times\exp(-j\omega\xi/U_c), \tag{1}$$

where U_c is the convective speed, ξ is the longitudinal separation, η is the lateral separation, and $S_{pp}(w)$ is the ordinary power spectrum of the pressure field.[6] Various investigators have identified the A and B functions based on empirical data. Although the Corcos model is simplistic, it is widely used in aeroacoustic models.

Strawderman developed the A and B functions for a zero pressure gradient flow with the exponential decays

$$A(\omega\xi/U_c)=e^{-\gamma_1|\omega\xi/U_c|},\qquad B(\omega\eta/U_c)=e^{-\gamma_3|\omega\eta/U_c|}, \tag{2}$$

where $\gamma_1=0.115$ and $\gamma_3=0.7$. The convective velocity (U_c) was assumed to be a constant given by $U_c=0.65U_\infty$ where U_∞ is the free stream velocity.[7]

A crude model for the ordinary power spectrum of the wall pressure was suggested by Skudrzyk and Haddle.[8] At the low normalized frequencies of concern, the ordinary power spectrum of the pressure field is approximately

$$S_{pp}=0.75\times10^{-5}\alpha^2\rho^2U_\infty^3\delta^*, \tag{3}$$

where $\alpha=3$ for air, ρ is the density of the fluid, and δ^* is the boundary layer displacement thickness.

By utilizing the Corcus model, the empirically determined exponential decays suggested by Strawderman, and the power spectra given by Skudrzyk and Haddle, the cross-spectral density of the wall pressure spectra was calculated. Using the cross-spectral density of the wall pressure spectra, the spectral density of the generalized force can be calculated. For this investigation, the cross terms of the spectral density of the generalized force are neglected. As discussed by Thomas and Nelson these cross terms are negligible when the main concern is the response at resonance for a system with low modal densities.[4] As shown by Thomas and Nelson the generalized force spectral density for the nth mode resulting from the turbulent flow field can be calculated as

$$S_{qq_n}(\omega)=\frac{1}{m^2}\int_{S_2}\int_{S_1}\Psi_{pq}(x_1,y_1)\Psi_{pq}(x_2,y_2)\times S_{ff}(\omega,x_2-x_1,y_2-y_1)\,dS_1\,dS_2, \tag{4}$$

where Ψ_{pq} are the orthonormal mode shapes of the plate, m is the mass per unit area of the plate, and S is the surface area of the plate. For a simply supported plate, the orthonormal mode shapes are

$$\Psi_{pq}(x,y)=\frac{2}{(ab)^{1/2}}\sin\left(\frac{p\pi x}{a}\right)\sin\left(\frac{q\pi y}{b}\right). \tag{5}$$

Integrating Eq. (4) gives

$$S_{qq_n} = \frac{4}{abm^2} A(B1+B2+B3+B4), \tag{6}$$

where

$$A = \frac{\gamma_3 \omega b/U_c}{[(\gamma_3\omega/U_c)^2+(q\pi/b)^2]} + \frac{2(q\pi/b)^2(1-\cos(q\pi)\exp(-\gamma_3\omega b/U_c))}{[(\gamma_3\omega/U_c)^2+(q\pi/b)^2]^2},$$

$$B1 = \frac{\omega a(\gamma_1+i)/2U_c}{(\omega(\gamma_1+i)/U_c)^2+(p\pi/a)^2},$$

$$B2 = \frac{\omega a(\gamma_1-i)/2U_c}{(\omega(\gamma_1-i)/U_c)^2+(p\pi/a)^2},$$

$$B3 = \frac{(p\pi/a)^2[1-\cos(p\pi)\exp(-\omega a(\gamma_1+i)/U_c)]}{[(\omega(\gamma_1+i)/U_c)^2+(p\pi/a)^2]^2},$$

and

$$B4 = \frac{(p\pi/a)^2[1-\cos(p\pi)\exp(-\omega a(\gamma_1-i)/U_c)]}{[(\omega(\gamma_1-i)/U_c)^2+(p\pi/a)^2]^2}.$$

To determine a state representation of the turbulent flow excitation, transfer functions between a white-noise input and the generalized forces must be found from the power spectra of the generalized forces developed using Eq. (6). This is accomplished using spectral factorization of the power spectra (Thomas and Nelson[4]). A transfer function having 20 numerator and denominator coefficients is used to represent the frequency information of the power spectra. Since the power spectra is real and positive, the odd ordered terms of the transfer function are zero. Furthermore, the transfer function is symmetric about the $j\omega$ axis. Thus the modal excitation filter is uniquely defined by the left half plane poles and zeros of the transfer function. In this way, a transfer function whose maximum size is tenth order is sufficient to model each generalized force. Finally, the modal excitation filter is converted from a transfer function representation to an observable canonical state space representation. A total of 296 states were used to describe the 50 modal excitation filters.

The power spectra calculated for the first mode [Eq. (4)] is shown in Fig. 2 as a solid line. The power spectra given by the spectral decomposition is shown as the dashed line. The difference between these two power spectra is less than 1 dB over the frequency range of interest. Each of the spectra for the 50 modes were fit within 10 dB across the frequency band from 10 to 10 000 Hz.

B. Structural response model

The structural response of the plate is obtained from the classical equation of motion for a thin damped flexible structure

$$D\nabla^4 u + c\frac{\partial u}{\partial t} + m\frac{\partial^2 u}{\partial t^2} = F(x,y,t), \tag{7}$$

where u is the normal displacement of the plate, D is the flexural rigidity, c is the viscous damping coefficient, m is the mass per unit area of the plate, and $F(x,y,t) = f(x,y)s(t)$ is the excitation pressure of the plate. Using the modal analysis method and assuming simply supported boundary conditions, Eq. (7) can be decoupled. The assumed solution is given by

$$u = \sum_{pq} \eta_{pq}\Psi_{pq}(x,y), \tag{8}$$

where η_{pq} are the modal displacements and Ψ_{pq} are the eigenfunctions. The modal equation can be written as

$$\ddot{\eta}_{pq} + 2\zeta\omega_{pq}\dot{\eta}_{pq} + \omega_{pq}^2\eta_{pq} = q_{pq}, \tag{9}$$

where ζ is the damping ratio and q_{pq} is the generalized modal force given by

$$q_{pq} = \frac{1}{m}\int_S \Psi_{pq}(x,y)f(x,y)dS. \tag{10}$$

For a point force excitation of the plate, the generalized modal forces are

$$q_{pq} = \frac{f}{m}\sin\left(\frac{p\pi x_0}{a}\right)\sin\left(\frac{q\pi y_0}{b}\right), \tag{11}$$

where x_0, y_0 are the coordinates of the excitation. For the turbulent boundary layer excitation, the output of the spectral factorization shaping filters are the generalized modal forces.

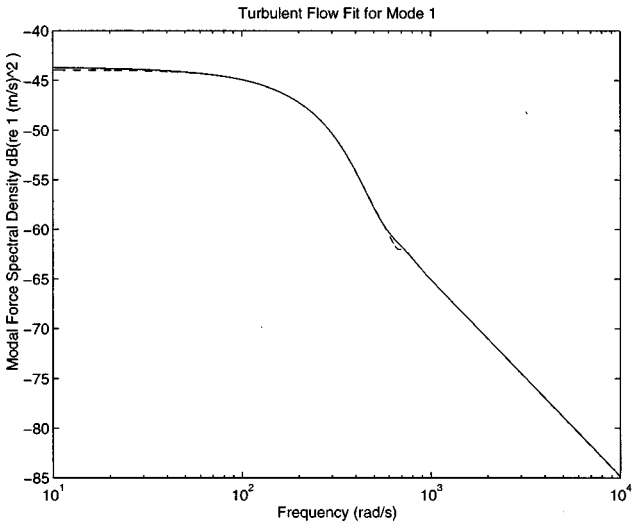


FIG. 2. Spectral decomposition for first mode.

C. Sound-pressure model

The third part of the analytical model is used to predict the sound-pressure level at a specified location resulting from the panel excitation. This is accomplished using the Rayleigh integral

$$p(\mathbf{R},t) = \frac{j\omega\rho_0}{2\pi}e^{j\omega t}\int_S \frac{v_n(\mathbf{r_s})e^{jk(|\mathbf{R}-\mathbf{r_s}|)}}{(|\mathbf{R}-\mathbf{r_s}|)}\,dS, \tag{12}$$

where ρ_0 is the density of air, k is the acoustic wave number, $\mathbf{r_s}$ is the position vector to the surface element, and $\mathbf{R}$ is the position vector to the control point. For a simply supported

TABLE I. System parameters.

Plate material	Aluminum
Plate dimensions (a,b,h)	(0.4, 0.2, 0.001) m
Modal damping ratio	0.01
Actuator location	$(a/2,b/2)$
Sensor locations	$(a/2,b/2)$
Controlled location	(2 m, 45°, 45°)
Fluid	Air
Air velocity U_∞	38 m/s (85 mile/h)
Convected velocity U_c	0.65 U_∞
γ_1	0.115
γ_3	0.7

plate, the normal velocity due to the pq mode can be written in terms of the modal acceleration as

$$v_{n_{pq}} = \frac{\ddot{\eta}_{pq}}{j\omega} \frac{2}{(ab)^{1/2}} \sin\left(\frac{p\pi x_0}{a}\right) \sin\left(\frac{q\pi y_0}{b}\right). \tag{13}$$

The frequency response functions $H_{pq}(\omega)$ for the component of the Rayleigh integral associated with each modal acceleration is

$$H_{pq}(\omega) = \frac{\rho_0 (ab)^{1/2} e^{-jkR}}{\pi^3 pq} \left[\frac{(-1)^q e^{j\beta} - 1}{(\beta/q\pi)^2 - 1}\right] \times \left[\frac{(-1)^p e^{j\alpha} - 1}{(\alpha/p\pi)^2 - 1}\right] e^{-j\beta/2} e^{-j\alpha/2}, \tag{14}$$

where

$$\alpha = ka \sin\theta \cos\phi, \qquad \beta = kb \sin\theta \sin\phi,$$

and $\mathbf{R}(R,\theta,\phi)$ specifies the location of the control point in spherical coordinates. The overall sound-pressure level at the control point is calculated using a modal summation of $H_{pq}(\omega)\,\ddot{\eta}_{pq}(\omega)$.

II. SYSTEM CONFIGURATION

The system to be controlled is a simply supported rectangular plate with a turbulent boundary layer excitation on one side. It is assumed that the plate radiates sound into an anechoic environment and the fluid loading of the air is negligible. The plate and fluid parameters are shown in Table I. A pictorial representation of the system is shown in Fig. 3.

The controller was configured for practical experimental implementation. A single-input single-output controller with a point force actuator is used. The feedback sensor is an accelerometer colocated with the control actuator in the center of the plate. The location of the sensor and actuator allow coupling only with the odd–odd plate modes. This is desirable since the odd–odd modes are the most efficient radiators of noise.[9]

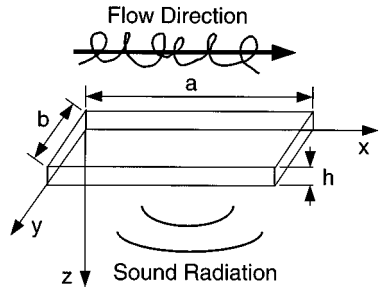


FIG. 3. Pictorial representation of the system.

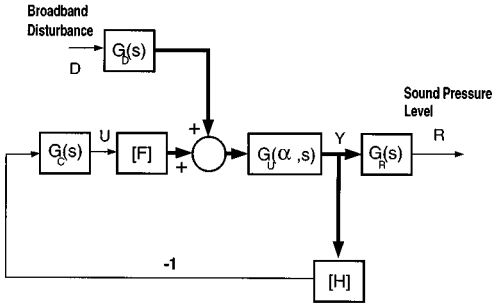


FIG. 4. Block diagram of the system.

The feedback structure is shown in Fig. 4. $G_D(s)$ represents the modal excitation due to the turbulent boundary layer given by the spectral factorization of Eq. (6), $G_U(\alpha,s)$ represents the dynamics of the plate, $G_R(s)$ represents the sound radiation to the control point due to the modal excitation given by Eq. (14), and $G_C(s)$ is the feedback controller. H and F are the modal participation coefficients associated with the sensor and actuator locations, respectively. The heavy lines indicate modal representations while the thin lines indicate measurable signals.

The system is subject to parametric uncertainty associated with variation in the natural frequencies and the damping ratio. The degree of parametric uncertainty for this investigation was chosen to be 1% variation in the natural frequencies and 5% variation in the damping ratio.

The performance goal is to reduce the sound pressure at a specified location to less than the desired sound-pressure level, P_{des}, subject to a control effort limitation, κ. However, the sound-pressure level is not used as feedback since for most applications it would be impractical to locate microphones at the locations where sound-pressure level reduction is desired. Furthermore, using the sound-pressure level as the feedback signal would restrict the controller bandwidth, which in turn, would limit closed-loop performance. This is a direct result of the phase lag associated with the delay of the sound propagation between the panel and the sensor location. As the distance between the sound-pressure feedback sensor and the radiating surface increases, so does the amount of delay.

Although the sound-pressure level is not used for feedback, it can be controlled indirectly when plate acceleration is used as feedback. Minimization of the plate acceleration does not guarantee that the sound-pressure level will be minimized. To effectively reduce sound radiation, the individual modal acceleration coefficients are weighted based on their sound-pressure level contribution at the output location. Radiating modes are heavily weighted and nonradiating modes are lightly weighted. In this way, the structural response of the plate is controlled such that the residual plate vibration occurs in modes which do not efficiently radiate sound and control effort is focused only on efficiently radiating modes.

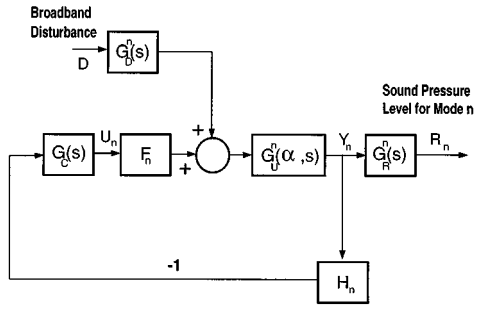


FIG. 5. Modal block diagram of the system.

III. FREQUENCY DOMAIN CONTROLLER DESIGN

For the frequency domain design methodology used in this investigation, the open-loop transfer function of the controller and system is shaped on a Nichols chart using classical frequency domain design tools. Frequency domain amplitude and phase bounds for disturbance rejection, stability, and control effort are generated on the Nichols chart and used to develop acceptable design regions for the design process. Loop shaping on the Nichols chart requires the designer to be involved in the controller design process. Because the disturbance rejection, control effort, and uncertainty bounds are plotted individually on the Nichols chart, the degree to which each bound influences the design is discernible and design sensitivity is apparent.

A. Developing frequency domain performance bounds

The sound-pressure level output performance bounds are developed on a mode by mode basis. Individual modal contributions to the sound-pressure level at the controlled position are restricted to be less than the desired sound-pressure level P_{des}. Thus bounds are generated based upon a system model associated with each individual mode. A block diagram of the system model for the nth mode is shown in Fig. 5. The modal acceleration $Y_n(s)$ for the nth mode can be written as

$$Y_n(\alpha,j\omega)=\frac{G_U^n(\alpha,j\omega)G_D^n(j\omega)}{1+G_U^n(\alpha,j\omega)F_nG_C(j\omega)H_n}\,D(j\omega). \quad (15)$$

Thus the sound pressure associated with the nth mode $[R_n(s)]$ can be written as

$$R_n(\alpha,j\omega)=\frac{G_R^n(j\omega)G_U^n(\alpha,j\omega)G_D^n(j\omega)}{1+G_U^n(\alpha,j\omega)F_nG_C(j\omega)H_n}\,D(j\omega). \quad (16)$$

The performance criteria is

$$\sum_{n=1}^{50}|R_n(j\omega)|\leqslant|P_{\mathrm{des}}(j\omega)|. \quad (17)$$

Using the nominal open-loop transfer function given by

$$L_0(\alpha_0,j\omega)=G_U(\alpha_0,j\omega)FG_C(j\omega)H \quad (18)$$

and substituting Eq. (16) into Eq. (17) gives the inequality

$$\left|\frac{1}{1+L_0(\alpha_0,j\omega)\mathscr{P}_n(\alpha,s)}\right|\leqslant\frac{P_{\mathrm{des}}(j\omega)}{|G_R^n(j\omega)G_U^n(\alpha,j\omega)G_D^n(j\omega)|}, \quad (19)$$

where

$$\mathscr{P}_n(\alpha,j\omega)=\frac{G_U^n(\alpha,j\omega)F_nH_n}{G_U(\alpha_0,j\omega)FH}. \quad (20)$$

Equating both sides of Eq. (19) gives an inverted M circle bound for each ω on the loop transfer function

$$L^n(\alpha,j\omega)=G_U^n(\alpha,j\omega)F_nG_C(j\omega)H_n. \quad (21)$$

For loop-shaping purposes, the bound is shifted in magnitude and phase by $\mathscr{P}_n(\alpha,j\omega)$ so that it becomes an amplitude bound on the nominal open-loop transfer function $L_0(\alpha_0,j\omega)$. The shifted inverted M circle bound is plotted on the Nichols chart to define the design region. To satisfy inequality Eq. (19), the magnitude of the nominal open-loop transfer function at ω for mode n and uncertainty α must be greater than the shifted inverted M circle.

The bounds developed using Eq. (19) are useful for the selection of the location for the control actuator and feedback sensor. Inappropriate sensor and actuator locations result in unachievable disturbance rejection bounds. For example, if

$$|P_{\mathrm{des}}(j\omega)|\leqslant|G_R^n(j\omega)G_u^n(\alpha,j\omega)G_D^n(j\omega)| \quad (22)$$

and either the sensor or the actuator is located on a node line, the disturbance rejection bound will have an infinite lower amplitude condition on the nominal open-loop transfer function.

Stability bounds are also shown on the Nichols chart. Stability is ensured by restricting the resonant peak of the closed-loop transfer function as

$$\left|\frac{L_0(\alpha_0,j\omega)\mathscr{P}(\alpha,s)}{1+L_0(\alpha_0,j\omega)\mathscr{P}(\alpha,s)}\right|\leqslant A, \quad (23)$$

where

$$\mathscr{P}(\alpha,j\omega)=\frac{G_U(\alpha,j\omega)}{G_U(\alpha_0,j\omega)} \quad (24)$$

and A is a constant. Equation (23) is an amplitude bound on the closed-loop transfer function which corresponds to an M circle bound on the open-loop transfer function. This M circle bound is shifted in magnitude and phase such that the amplitude bound corresponds to the nominal open-loop transfer function $L_0(\alpha_0,j\omega)$. A satisfactory $L_0(\alpha_0,j\omega)$ has magnitude and phase characteristics that are outside the M circle bound.

Finally, control effort bounds are developed and shown on the Nichols chart. These amplitude bounds are used to ensure that the controller does not saturate the actuator. The actuator constraint is

$$|u(t)|\leqslant\kappa, \quad (25)$$

where $u(t)$ is the control effort. The closed-loop transfer function relating the disturbance $D(s)$ to the control effort $U(s)$ is

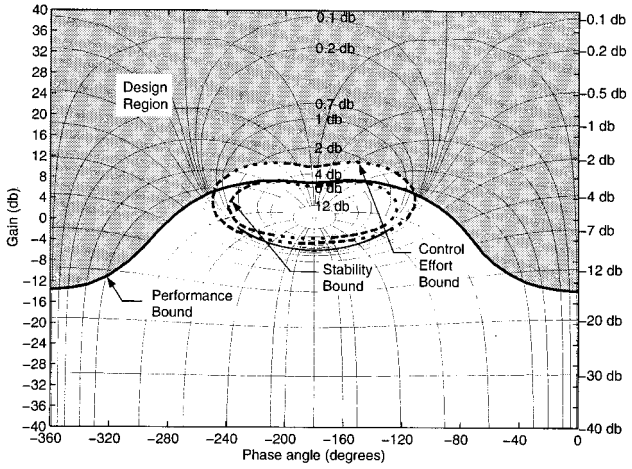


FIG. 6. Composite bound for $\omega = 1258$ rad/s.

$$\frac{U(\alpha,j\omega)}{D(j\omega)} = \frac{-G_C(j\omega)HG_U(\alpha,j\omega)G_D(j\omega)}{1+G_U(\alpha,j\omega)FG_C(j\omega)H} \tag{26}$$

$$= \left(\frac{-G_D(j\omega)}{F}\right)\left(\frac{L_0(\alpha_0,j\omega)\mathscr{P}(\alpha,s)}{1+L_0(\alpha_0,j\omega)\mathscr{P}(\alpha,s)}\right). \tag{27}$$

For the turbulent flow, the disturbance input to the shaping filters is broadband, $D(j\omega)=1$. Franchek[10] and Jayasuriya[11] have shown that a time-domain control effort response can be directly limited by a frequency domain bound. Such a frequency domain bound for control effort is developed from Eq. (27) and is given by

$$\left|\frac{L_0(\alpha_0,j\omega)\mathscr{P}(\alpha,s)}{1+L_0(\alpha_0,j\omega)\mathscr{P}(\alpha,s)}\right| \leqslant \left|\frac{\kappa F}{G_D(j\omega)}\right|. \tag{28}$$

Equation (28) represents an M circle bound on the open-loop transfer function. The frequency characteristics of the open-loop transfer function must be outside the M circle.

B. Bound generation

Using Eq. (19), performance bounds were generated from each of the first 50 plate modes with the desired pressure chosen as $P_{\text{des}}=56$ dB. Stability bounds were generated using Eq. (23) with an M circle of 6 dB. This provides a minimum gain margin of 3.5 dB and phase margin of 30°. Using Eq. (28), control effort bounds using $\kappa = 14\ \mu$N were generated. All three performance bound types were generated using a gridding of the parametric uncertainty of 1% natural frequency variation and 5% damping ratio variation.

The disturbance rejection, stability, and control effort bounds were combined to form the acceptable design region for various individual frequencies. An example set of bounds developed at the natural frequency of the third mode ($\omega = 1258$ rad/s) is shown in Fig. 6. The nominal open-loop transfer function at $\omega = 1258$ rad/s must be within the shaded design region to satisfy inequalities Eqs. (19), (23), and (28). The frequencies for which the design region was most restrictive were used in the loop shaping process.

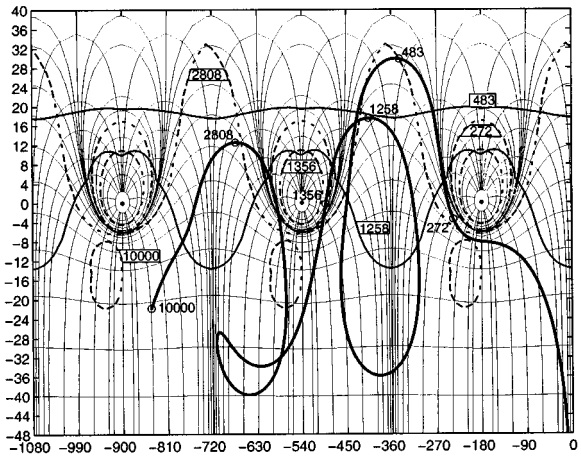


FIG. 7. Performance bounds for frequency domain controller design.

C. Loop shaping the controller

The performance bounds developed from the frequency domain amplitude inequalities impose amplitude and phase conditions on the open-loop transfer function to be designed. These bounds are transformed into bounds confining the nominal open-loop transfer function using the inverse normalized plant template $\mathscr{P}_n^{-1}(\alpha,j\omega)$. For this design, the performance bounds requiring low gain on the sensitivity transfer function occur at the natural frequencies of the odd-odd modes of the plate. This was anticipated since the odd–odd modes are the most efficient sound radiating modes.

From the stability requirements based on the Nyquist stability criterion, the phase lag due to the plate dynamics requires the gain of the open-loop transfer function to be below unity for phases at and near odd multiples of $-180°$. The frequency domain bounds for control effort impose a low gain condition on the complementary sensitivity transfer function which was most restrictive at the off-resonance frequencies. For a lightly damped structure, the higher-order modes contribute much more significantly to the frequency response at off-resonant frequencies than at resonant frequencies. Therefore, the fact that these modes are unmodeled contributes significantly to the uncertainty at the off-resonant frequencies. This causes the control effort and stability bounds to be more restrictive at off-resonant frequencies.

The controller is designed by augmenting the open-loop transfer function with additional poles and zeros such that the amplitude and phase characteristics satisfy the frequency domain bounds. Shown in Fig. 7 is the open-loop transfer function and a few key frequency domain bounds on an extended Nichols chart. The open-loop transfer function is shown as a heavy solid line. The frequencies associated with each bound and the corresponding frequency on the open-loop transfer function are labeled. Lower bounds are solid and upper bounds are dashed. Due to the multiple modes of the plate, the phase lag associated with the plant dynamics is less than $-180°$.

The performance bounds require the gain of the open-loop transfer function to be greater than unity for phases less than $-180°$ lag. To achieve this performance requirement,

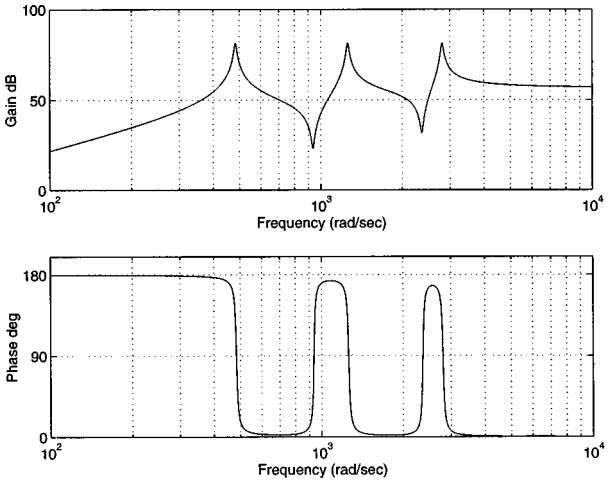


FIG. 8. Bode plot of the nominal plant $[FG_u(\alpha_0,s)H]$.

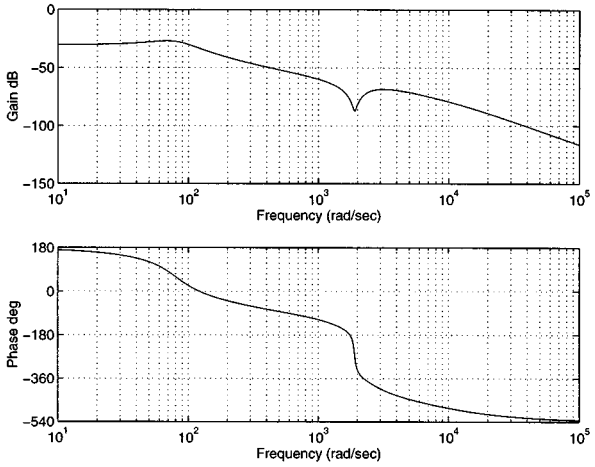


FIG. 9. Bode plot of $G_C(s)$.

the controller could be used to add phase lead to compensate for the excessive phase lag due to the plant dynamics. However, this type of feedback controller would be susceptible to sensor noise. Alternatively, the feedback controller could be used to add additional phase lag such that the performance requirements are satisfied while also meeting the Nyquist stability criterion. In this investigation, nonminimum phase zeros are utilized for this purpose. It is important to note that the natural dynamics of the system (Fig. 8) produces the amplitude characteristics required by the performance bounds. With robust frequency domain controller design it is possible to utilize these natural dynamics in the controller design thereby reducing the required controller gain and bandwidth.

The controller design begins as a negative gain. This shifts the phase by $-180°$. Next, a complex pole at $\omega = 80$ rad/s, $\zeta = 0.4$, and a nonminimum phase zero at $\omega = 150$ rad/s are used to add phase lag so that the first and third natural frequencies ($\omega = 483$ and 1258 rad/s) are centered between the stability points at $-180°$ and $-540°$. The amplitude roll-up associated with the nonminimum phase zero is attenuated by the amplitude roll-off of the complex poles. Therefore, the desired phase characteristic is achieved without decreasing the natural plant gain at the modal resonances. The upper bounds at $\omega = 272$ and 1356 rad/s are met by limiting the controller gain. The lower bounds at $\omega = 483$ and 1258 rad/s are easily achieved. A nonminimum phase complex conjugate pair of zeros at $\omega = 1900$ rad/s, $\zeta = 0.03$ was then used to add phase lag to place the resonance peak corresponding to the eighth mode ($\omega = 2808$ rad/s) between the stability points at $-540°$ and $-900°$. These zeros also add enough gain near this frequency to push the amplitude at $\omega = 2808$ rad/s to the upper bound limit. Since the performance bounds for higher frequencies (i.e., $\omega = 10\,000$ rad/s) are small inverted M circles, these bounds are easy to satisfy. Therefore, the controller can be used to roll-off the open-loop gain without violating the stability, performance, or control effort bounds. To roll-off the gain of the controller at high frequencies, complex poles were added at $\omega = 2200$ rad/s with $\zeta = 0.5$. These poles also add phase lag to help space the eighth modal resonance. An additional pole at $\omega = 9000$ rad/s is used to increase this amplitude roll-off. The resulting controller is a fifth-order controller with a dc gain of -30 dB,

$$G(s) = \frac{-0.03[-s/150+1][s^2/1900^2 - 2(0.03)s/1900+1]}{[s^2/80^2+2(0.4)s/80+1][s^2/2200^2+2(0.5)s/2200+1][s/9000+1]}. \tag{29}$$

By utilizing the natural dynamics of the plate, the controller order and gain were minimized. The bode plot of the controller is shown in Fig. 9.

IV. CONTROLLER PERFORMANCE

Using the controller designed in the previous section, the sound-pressure level of the control point is calculated for the flow noise disturbance model. The analytical model was used to evaluate the performance. The uncontrolled and controlled sound-pressure level responses are both shown. Additionally, the time history of the control effort and the uncontrolled and controlled sensor acceleration are given for the turbulent flow disturbance.

The uncontrolled and controlled sound-pressure level responses are shown in Fig. 10. The sound-pressure level response for the flow disturbance input is reduced by approximately 16 and 8 dB at the first and third resonances, respectively. Additional reductions are achieved at higher odd–odd resonance frequencies. The desired performance

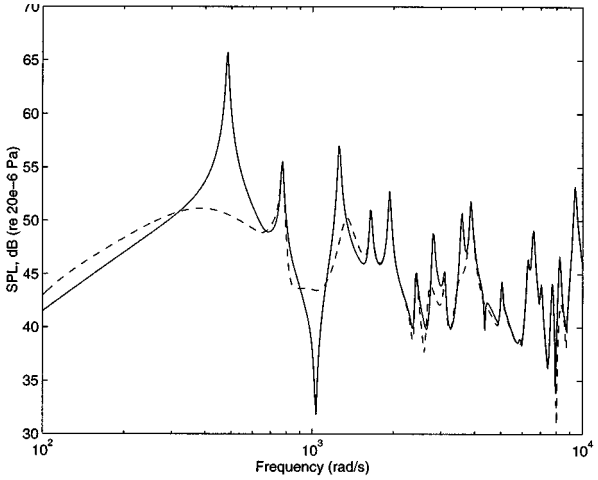


FIG. 10. Sound-pressure level for the turbulent flow excitation, uncontrolled —; controlled --.

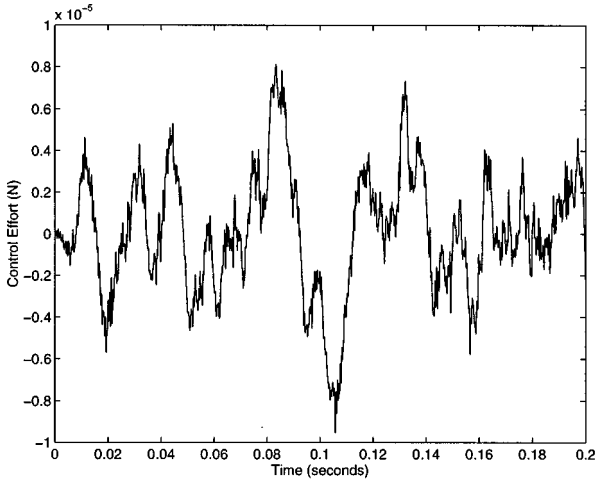


FIG. 12. Control effort response to flow disturbance.

(P_{des}=56 dB) was achieved at all frequencies. In fact substantially lower sound-pressure levels were achieved at frequencies where the nominal open loop transfer function greatly exceeded the minimum performance bounds.

The time history of the acceleration at the sensor location is shown in Fig. 11. The rms value for the controlled case is 3.8×10^{-3} m/s^2 while that of the uncontrolled case is 6.1×10^{-3} m/s^2. The control effort is limited to less than 10 μN and is shown in Fig. 12. This is less than the control limit bound of 14 μN. Despite the low controller order, dc gain, and control effort, the performance of the controller is significant. Furthermore, controller implementation is feasible unlike previous active controllers for the flow-induced sound radiation problem.

V. CONCLUSION

This work has shown that a practical feedback controller implementation for the flow induced structural radiation of sound problem is possible. A robust control methodology was successfully used in the controller design. The controller design incorporated uncertainty to ensure robust performance. Significant sound-pressure level reductions were obtained at the odd–odd resonances.

The superiority of a frequency domain controller approach for this problem has been demonstrated. The performance goals for this technique have been formulated to ensure that a specific modal sound-pressure level is achieved for a turbulent flow disturbance. By representing the performance, stability, and control effort criteria as bounds on the Nichols chart, their impact on the controller design can easily be visualized. Furthermore, the method achieved good performance using a small controller dc gain and order. This result is obtained as a direct consequence of the use of nonminimum phase zeros to add controller lag such that the beneficial plant characteristics are utilized.

Subsequent work will extend the controller design methodology to multiple-input multiple-output controllers. In this way control of the even modes can be achieved. An experimental implementation of control of turbulent induced structural sound radiation will also be undertaken.

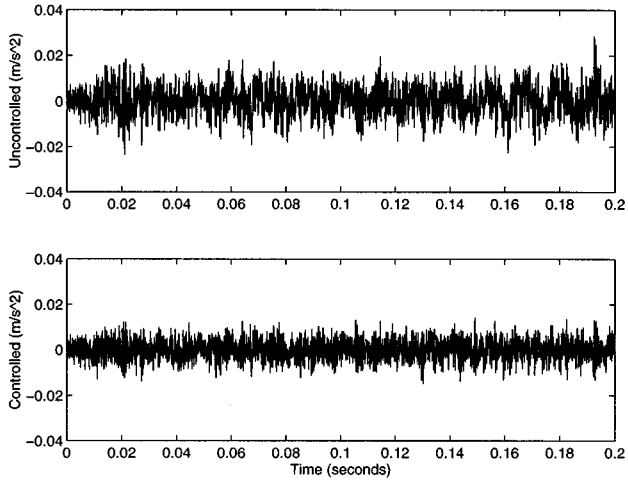


FIG. 11. Sensor acceleration for flow disturbance.

[1] X. H. Yang, J. van Niekerk, K. S. Parwani, K. Hendrick, A. Packard, and B. Tongue, ''Acoustical response tailoring in reverberant enclosures using feedback control,'' ASME Active Control of Noise and Vibration **38**, 7–14 (1992).
[2] X. H. Yang, J. van Niekerk, K. S. Parwani, A. Packard, and B. Tongue, ''Attenuation of structurally generated interior noise through active control,'' in *Proceedings of the 1993 American Control Conference* (American Automatic Control Council, 1993), pp. 1–7.
[3] O. Fluder and R. Kashani, ''Robust control of structure-borne noise using h_∞ methods,'' ASME Active Control of Noise and Vibration **38**, 191–204 (1992).
[4] D. R. Thomas and P. A. Nelson, ''On the use of feedback control of sound radiation from a plate excited by a turbulent boundary layer,'' J. Acoust. Soc. Am. **98**, 2651–2662 (1995).
[5] W. A. Strawderman, ''Turbulence-induced plate vibrations: And evaluation of finite- and infinite-plate models,'' J. Acoust. Soc. Am. **46**, 1294–1307 (1969).
[6] G. M. Corcos, ''Resolution of pressure in turbulence,'' J. Acoust. Soc. Am. **35**, 192–199 (1963).
[7] W. A. Strawderman and R. S. Brand, ''Turbulent-flow-excited vibration of a simply supported, rectangular flat plate,'' J. Acoust. Soc. Am. **45**, 177–192 (1968).

[8] E. J. Skudrzyk and G. P. Haddle, ''Noise production in a turbulent boundary layer by smooth and rough surfaces,'' J. Acoust. Soc. Am. **32**, 19–34 (1960).
[9] F. Fahy, *Sound and Structural Vibration Radiation, Transmission and Response* (Academic, New York, 1985).
[10] M. Franchek, ''Selecting the performance weights for the μ and h_∞ synthesis methods for siso regulating,'' ASME J. Dynamic Syst., Measurement, Control **118**, 126–131 (1996).
[11] S. Jayasuriya, ''Qft type design for maximizing tolerable disturbances in structured uncertain systems,'' Int. J. Control **56**, 67–85 (1992).

Effects of high-sided vehicles on the performance of noise barriers

D. C. Hothersall and S. A. Tomlinson
Department of Civil and Environmental Engineering, University of Bradford, Bradford BD7 1DP, United Kingdom

(Received 28 February 1996; revised 21 October 1996; accepted 16 March 1997)

The attenuation produced by a given barrier can be maximized by placing it as close to the noise source or receiver as possible. This fact and the availability of land often means that barriers are situated close to the edge of the road. In this position the performance of the barrier can be degraded by the effects of multiple reflections of sound between the sides of the vehicles and the barrier. This paper reports the results of calculations, carried out using a two-dimensional boundary element model, of the attenuation produced by various configurations of barrier adjacent to a traffic stream. Spectra of insertion loss and the mean broadband insertion loss in the shadow region are considered. The effects of the height of the vehicle, the distribution of absorbing material on the traffic-facing side of the barrier, and the slope of this face are considered quantitatively in terms of the peak noise level. It is concluded that serious degradation of the performance of the barrier can occur for some situations, especially when the side of the vehicle protrudes above the barrier. The results indicate the extent to which the problem can be overcome by the use of absorbing barriers or by sloping barriers of suitable design. [S0001-4966(97)04907-2]

PACS numbers: 43.50.Gf, 43.50.Lj [GAD]

INTRODUCTION

Accurate prediction of the noise from vehicles is important as a tool in the design and planning of transport schemes and in assessments for compensation. Such prediction schemes exist in most countries. In considering the effects of noise barriers the attenuation produced is usually calculated by a path difference approach, as in Ref. 1. In this case the difference between the length of the direct path from source to receiver and the path via the top of the barrier is related to the attenuation. The attenuation increases with the height of the barrier. The method also predicts that the attenuation is optimized by placing the barrier as close to the source as is practicable. This consideration and also the scarcity of land in many urban situations has sometimes resulted in the siting of noise barriers very close to roads or railways. In these conditions it is possible that strong reflections of the sound between the sides of the vehicles and the barrier may reduce the efficiency of the barrier from that predicted by a simple path difference approach. This effect has been discussed by several authors (e.g., Ref. 2) but there is no clear agreement of the extent of the effect, and little quantitative information is available. One method of modeling this phenomenon is to extend the path difference model using multiple images of the source, but in the basic model the phase relationships between the contributions from the source and each image are not considered and the reflection and diffraction mechanisms are simplified. A second approach is to develop a numerical wave solution for the sound field. In this paper results are reported for two-dimensional numerical modeling of these conditions using a boundary integral equation method. The effects of the proximity and the height of the vehicle on the performance of a plane wall barrier are reported (see also Refs. 3 and 4). The use of acoustically absorbing surfaces and sloping sections on the traffic-facing side of the screens to overcome the problem is considered. Although these treatments have been used in practical cases, published results accurately quantifying the efficiency of these measures are scarce.

I. NUMERICAL MODEL

The numerical model is applied to the system shown in Fig. 1. The body of the vehicle is represented as a rectangle elevated above the flat ground surface with the source offset 0.03 m horizontally from the lower corner nearest the barrier. The three-dimensional representation is equivalent to an infinite coherent line source with an infinitely long barrier and vehicle parallel to the source and of uniform cross section and surface covering along their length.

In general terms let $p(\mathbf{r},\mathbf{r}_0)$ represent the acoustic pressure at $\mathbf{r} = (x,y)$ with the source at $\mathbf{r}_0 = (x_0,y_0)$. The x axis lies in the ground plane and the y axis is perpendicular to this plane, into the propagating medium. Let γ define the surfaces above the ground plane and it is assumed that these are locally reacting with normalized surface admittance $\beta(\mathbf{r}_s)$, where $\mathbf{r}_s$ is the position vector of a point on γ. A reformulation of the Helmholtz equation as a boundary integral equation in which the integral is taken over the surfaces of the barrier and vehicle is solved at a single frequency by a boundary element approach.[5,6] To carry this out the surfaces of the vehicle and the barrier, γ, are divided into a number of straight line elements $\gamma_1, \gamma_2, \gamma_n, \cdots, \gamma_N$. The acoustic pressure, $p(\mathbf{r}_n,\mathbf{r}_0)$ is assumed constant over each element and is calculated at the mid-point. $\mathbf{r}_n$ is the position vector of the mid-point of the nth element. As a result of this approximation the integral equation reduces to ($e^{-i\omega t}$ time dependence)

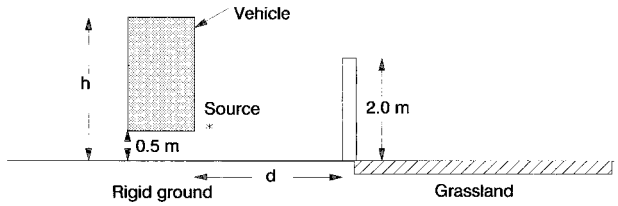


FIG. 1. Configuration of the two-dimensional model.

$$\varepsilon(\mathbf{r})p(\mathbf{r},\mathbf{r}_0)=G_{\beta_c}(\mathbf{r}_0,\mathbf{r})+\sum_{n=1}^{N} p(\mathbf{r}_n,\mathbf{r}_0)\int_{\gamma_n}\left(\frac{\partial G_{\beta_c}(\mathbf{r}_s,\mathbf{r})}{\partial n(\mathbf{r}_s)}\right.$$

$$\left.-ik\beta(\mathbf{r}_s)G_{\beta_c}(\mathbf{r}_s,\mathbf{r})\right)ds(\mathbf{r}_s), \tag{1}$$

where $ds(\mathbf{r}_s)$ denotes the length of an element of γ at $\mathbf{r}_s$, $\partial/\partial n(\mathbf{r}_s)$ denotes the partial derivative in the direction of the normal to γ at $\mathbf{r}_s$ directed out of the propagating medium, and k is the wave number. The integral is evaluated using a product mid-point rule approximation described in Ref. 7. Here, $\varepsilon(\mathbf{r})$ has a value of unity when $\mathbf{r}$ lies anywhere in the propagating medium except on γ; $\varepsilon(\mathbf{r})=1/2$ if $\mathbf{r}$ is a point on γ which is not a corner point. If $\mathbf{r}$ is a corner point then $\varepsilon(\mathbf{r})=\Omega/(2\pi)$, where Ω is the angle in the medium subtended by the two tangents to the boundary at $\mathbf{r}$. β_c is the normalized surface admittance of the ground plane which is assumed to be constant, and $G_{\beta_c}(\mathbf{r},\mathbf{r}_0)$ denotes the acoustic pressure at $\mathbf{r}$ due to the source at $\mathbf{r}_0$ in the presence of the ground plane only. $G_{\beta_c}(\mathbf{r},\mathbf{r}_0)$ can be written as[8]

$$G_{\beta_c}(\mathbf{r},\mathbf{r}_0)=-\frac{i}{4}\{H_0^{(1)}(k|\mathbf{r}_0-\mathbf{r}|)+H_0^{(1)}(k|\mathbf{r}_0'-\mathbf{r}|)\}$$

$$+P_{\beta_c}(\mathbf{r},\mathbf{r}_0), \tag{2}$$

where $\mathbf{r}_0'=(x_0,-y_0)$ is the image of the source in the ground plane, $H_0^{(1)}$ is the Hankel function of the first kind of order zero, and

$$P_{\beta_c}(\mathbf{r},\mathbf{r}_0)$$

$$=0,\quad \beta_c=0,$$

$$=\frac{i\beta_c}{2\pi}\int_{-\infty}^{\infty}\frac{\exp(ik((y+y_0)(1-s^2)^{1/2}-(x-x_0)s))ds}{(1-s)^{1/2}((1-s^2)^{1/2}+\beta_c)},$$

$$\mathrm{Re}\,\beta_c>0, \tag{3}$$

with $\mathrm{Re}[(1-s^2)^{1/2}],\mathrm{Im}[(1-s^2)^{1/2}]\geqslant 0$. Efficient and accurate methods of calculating P_{β_c} and $\partial P_{\beta_c}/\partial n$ when $\mathrm{Re}\,\beta_c>0$ are detailed in Ref. 8. The boundary element method then follows two steps. By setting $\mathbf{r}$ equal to $\mathbf{r}_m$ for $m=1,2,\ldots,N$ in Eq. (1) a set of N linear equations in the unknowns $p(\mathbf{r}_1,\mathbf{r}_0),p(\mathbf{r}_2,\mathbf{r}_0),\ldots,p(\mathbf{r}_n,\mathbf{r}_0),\ldots,p(\mathbf{r}_N,\mathbf{r}_0)$ is obtained. After solving these equations, defining the pressure at the mid-point of each element, Eq. (1) can be used as a formula to calculate the pressure at a receiver point in the propagating medium.

Virtually any shape of barrier cross section and distribution of surface acoustic properties over the section can be considered using this method. The rigid section of road surface is introduced by defining a section of γ in the uniform, finite admittance ground plane. The standard barrier used throughout is 0.2 m in width and 2.0 m in height. Barriers of this height are used in many countries and the height was selected so that the effects of the vehicle body extending above the barrier could be investigated. The source position is approximately representative of noise from tires, engine, power train, and a low-mounted muffler. It would be necessary to introduce a separate source to represent a high-mounted muffler and this has not been considered. For the barrier height used a source in this elevated position would be directly visible at all the receiver points. Consider the case where the barrier is sufficiently high to shield the whole vehicle. A muffler protruding above the top of the vehicle is outside the region defined by the parallel surfaces of the barrier and the side of the vehicle. In these conditions multiple reflections between these surfaces will be of little significance.

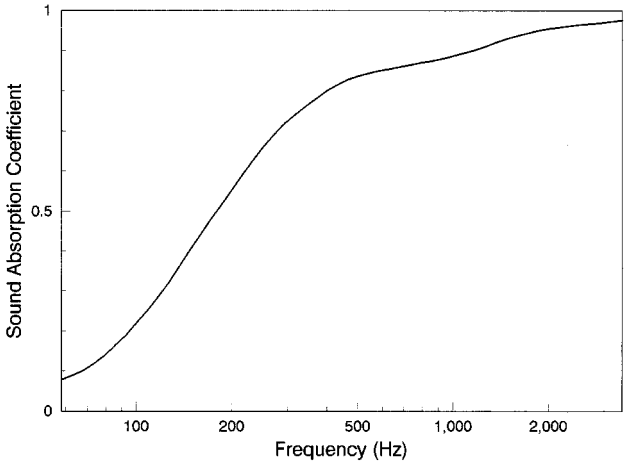


FIG. 2. Normal incidence sound absorption coefficient calculated using the Delany and Bazley model for admittance of a fibrous absorbing material with flow resistivity of 20 000 N s m^{-4} and a layer depth of 0.1 m.

The accuracy of the model has been investigated in Ref. 6 by direct comparison with an exact analytical solution of an extended barrier problem. At frequencies below 2500 Hz the differences were always less than 0.5 dB. For the results from 2500 to 3150 Hz larger discrepancies were observed, of the order of 5 dB.

The surface admittance of the absorbent surfaces is calculated according to the equations of Delany and Bazley[9] for fibrous materials. The flat ground on the side of the barrier away from the road is assumed to be grassland, having a flow resistivity of 200 000 N s m^{-4} and a layer depth of 0.1 m. Where the use of absorbing material is specified on a barrier, a flow resistivity of 20 000 N s m^{-4} and a layer depth of 0.1 m is used. This results in a surface impedance similar to that of a layer of mineral wool or a similar strong absorber and a spectrum of the normal incidence sound absorption coefficient calculated using these parameters is shown in Fig. 2.

The acoustic pressure is calculated at 1/9 octave center frequencies between 58 and 3415 Hz for nine receiver positions at heights of 0, 1.5, and 3.0 m above the ground and at distances of 20, 50, and 100 m from the center line of the

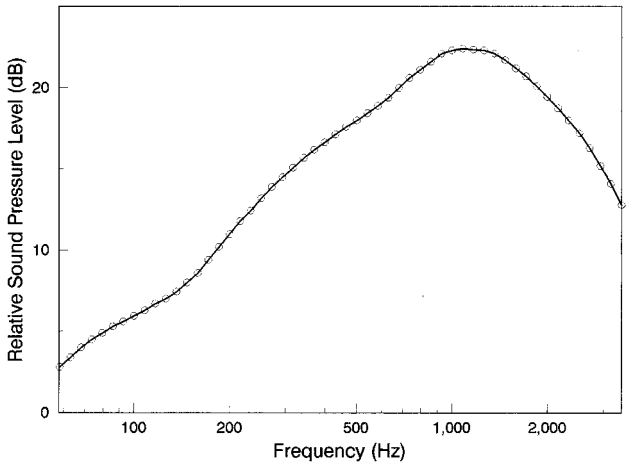


FIG. 3. 1/9th-octave relative sound-pressure level spectrum characteristic of A-weighted roadside traffic noise.

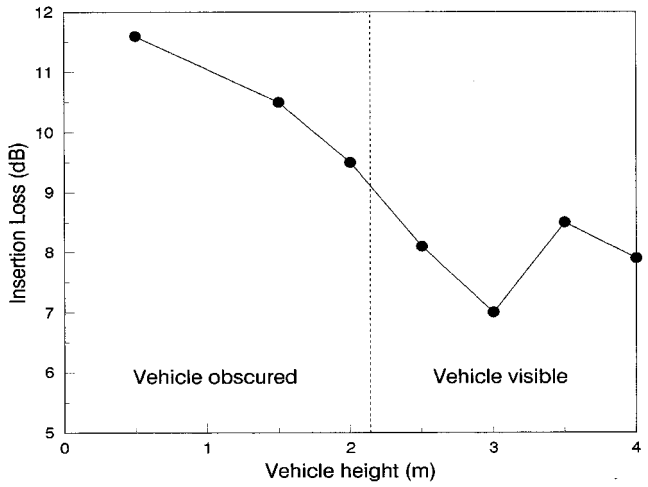


FIG. 4. Broadband insertion loss at a receiver position 20 m behind the barrier and 1.5 m above the ground as a function of vehicle height. The barrier has vertical rigid sides and is 2.0 m high. The vehicle is 5.5 m from the barrier.

barrier. It is important to carry out the calculations at these relatively small frequency intervals in order to resolve the rapid changes in the spectrum associated with interference effects arising from the multiple reflections between the vehicle and the barrier. The insertion loss at each frequency is calculated as

$$\mathrm{IL} = 20 \log_{10} |P_g / P_b| \quad \mathrm{dB}, \tag{4}$$

where P_g is the sound pressure at the receiver when source, vehicle, and ground are present, and P_b is the sound pressure at the receiver for the same conditions with the barrier included.

In order to calculate the broadband insertion loss details of the source spectrum shape are required. The spectrum used is characteristic of roadside A-weighted road traffic noise and is shown in Fig. 3. It is based on a 1/3 octave spectrum given in Ref. 10, which is derived from measurements in various European countries and reviewed within the noise barrier standards committee CEN/TC226/WG6/TG1. The results for the acoustic pressure at the 1/9 octave center frequencies calculated using the boundary element method were weighted using the traffic noise spectrum and then combined. The broadband insertion loss is then defined in the same way as in Eq. (4). Although a variation in the absolute values of insertion loss was observed at the different receiver positions, the *changes* observed as the vehicle and barrier conditions were altered were similar for each receiver position. The arithmetic mean of the broadband insertion loss values over the six receiver positions above the ground can thus be considered as a general measure of the relative efficiency of the barrier systems in the shadow region and this figure is defined as the *mean insertion loss*. The use of this indicator is discussed further in Ref. 6.

The computing resources necessary to carry out the calculations primarily depend on the length of γ and the frequency. The results reported here were calculated on a networked workstation system and the CPU time for each configuration modeled was between approximately 100 and 600 s. It would be desirable to improve the simulation by the use of a three-dimensional boundary element model. With severe truncation of such a model in the third dimension (along the roadway) and assuming only rigid boundary conditions, it is possible to calculate a solution for very low frequencies. For more complex site conditions and higher frequencies the computing resources required are prohibitive.

II. RESULTS

A. Effect of vehicle height

In Fig. 4 the broadband insertion loss at a receiver position 20 m behind the barrier and 1.5 m above the ground is plotted as a function of the height of the vehicle. The barrier has rigid surfaces. The distance from the barrier to the side of the vehicle is 5.5 m and this distance is used as a standard throughout, except in Sec. II B. From this receiver position the vehicle becomes visible over the barrier at a height of 2.14 m. There is a reduction in insertion loss as the vehicle height increases from 1.5 to 2.5 m. The insertion loss remains approximately constant when vehicle height is greater than 2.5 m.

The mean insertion loss over six receiver positions is plotted as a function of vehicle height in Fig. 5. The solid line is for a barrier with rigid surfaces. When $h = 0.5$ m the side of the vehicle is not present. The mean insertion loss reduces as the height of the vehicle is increased. A value of $h = 1.5$ m corresponds approximately to the dimension of an automobile. The vehicle becomes visible at the different receiver positions at heights of between 1.7 and 2.2 m. When the height of the vehicle is greater than that of the barrier, there is a 5-dB reduction in efficiency compared with the result when the body is not included in the calculation.

Spectra of insertion loss at a receiver position of 50 m behind the barrier and in the ground surface are shown in Fig. 6. The solid line is for the source at 5.47 m from the barrier in the absence of the vehicle body. The oscillations in the spectrum result from interference between the direct rays and the ground reflected rays on the source and receiver sides of the barrier. The dashed line is for similar conditions but with a vehicle height of 4.0 m. The overall trend of the curve

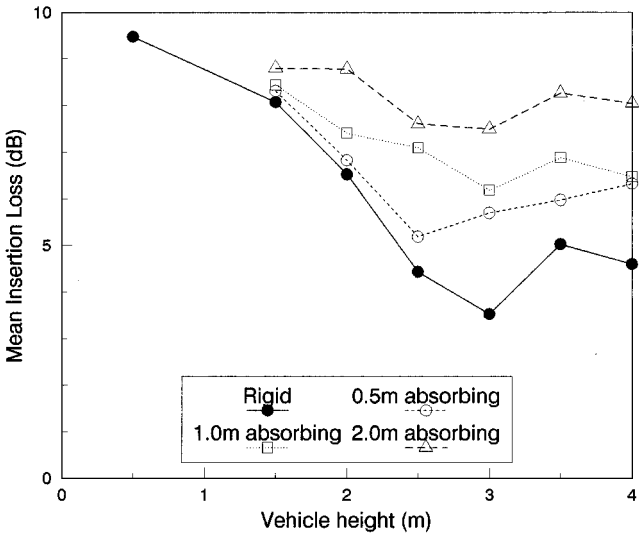


FIG. 5. Mean insertion loss as a function of vehicle height for various absorbing surface treatments of the road-facing side of the barrier. The barrier has vertical sides and is 2.0 m high. The vehicle is 5.5 m from the barrier.

obtained without the vehicle body can still be observed but in addition there are rapid oscillations in this spectrum due to the path differences introduced by strong multiple reflections between the side of the vehicle and the rigid barrier.

B. Effect of the separation of vehicle and barrier

A series of predictions were calculated in which the distance (d) between the side of the vehicle and the barrier was changed. The height of the vehicle remained constant at 4.0 m. For d = 2.0, 5.5, 18.5, and 22.0 m the mean insertion losses were 4.6, 4.6, 4.2, and 4.0 dB, respectively. These results are quite similar and do not show the expected trend of increasing insertion loss with the separation of the vehicle and the barrier. This may be a consequence of the geometrical accuracy of the model. In practice the sides of the vehicles and the barrier would not be exactly parallel and the vehicles would be of finite length. This would be likely to reduce the multiple reflection effects at the greater distances.

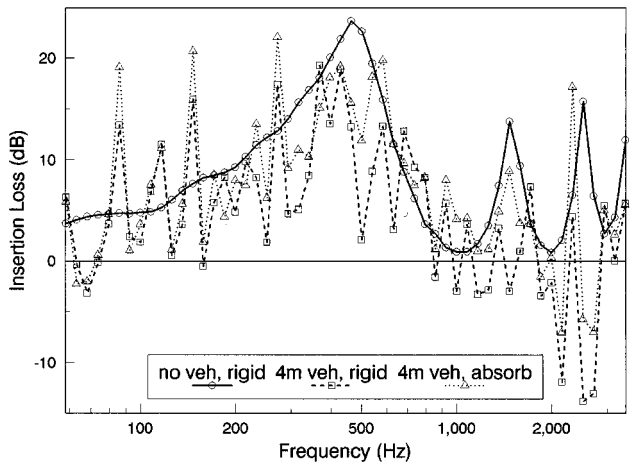


FIG. 6. The 1/9-octave spectra of insertion loss at a receiver position 50.0 m from the barrier, in the ground surface; (a) in the absence of the vehicle body, with the source 5.47 m from the barrier which has a rigid surface, (b) for a vehicle of height 4.0 m at 5.5 m from a barrier with a rigid surface, and (c) for a vehicle of height 4.0 m at 5.5 m from a barrier with an absorbing surface.

C. Effect of an absorbing surface

In order to reduce the effects of multiple reflections between the side of the vehicle and the barrier various distributions of absorbing material were placed on the traffic-facing side of the barrier. Details of the assumed characteristics of the material are given in Sec. I. Results for the mean insertion loss in conditions where the absorber covers the upper 0.5 m, the upper 1.0 m, and the whole barrier surface facing the traffic are shown in Fig. 5 as a function of vehicle height. When the top 0.5 m is absorbing an increase of between 1 and 2 dB in the mean insertion loss is observed over the results for the rigid surface in the region where the vehicle is visible over the barrier. When the absorbing strip is extended a further increase is observed. The results when the whole surface is treated are similar to that for a rigid barrier with an automobile present (height 1.5 m) for all the vehicle heights. This value is approximately 1 dB lower than the mean insertion loss in the absence of a vehicle body; corresponding to the result for a height of 0.5 m in Fig. 5. These results are for a system in which the sound source is assumed to be close to the ground. If sources also exist near the top of the vehicle (e.g., vertical exhaust muffler systems) the results would be significantly modified.

In Fig. 6 the effect on the insertion loss spectrum when the whole of the traffic-facing side of the barrier is absorbing can be observed. The oscillations in the spectrum resulting from the multiple reflections between the rigid surfaces are reduced, particularly at the higher frequencies.

D. Effect of sloping surfaces

A method which has been proposed to reduce the effects of multiple reflections in outdoor sound propagation is to slope the reflecting surfaces from the vertical. The sound energy is thus directed upward either directly, if the barrier slopes away from the source, or after an intermediate reflection at the ground if the surface slopes toward the source. The argument assumes that the barrier surface acts as a specular reflector of the sound. This will be the dominant process for wavelengths less than the dimensions of the surface. For longer wavelengths a scattering effect will occur. With periodic sloping structures diffraction grating effects are also possible.

Figure 7 shows three cross sections which were considered. The barrier comprises one, two, or four panels with rigid surfaces and an outward sloping face at an angle of θ from the vertical. It has been suggested that θ=10° is a suitable angle to adopt for parallel barriers in many situations to reduce reflected sound to an insignificant level at typical source and receiver positions.[11,12] In Fig. 8 the mean insertion loss of these designs for θ=10° is plotted as a function of the height of the vehicle and compared with results for the vertical rigid screen of the same height. The sloping faces produce some marginal improvement in the performance of the barrier for high sided vehicles, of the order of 1 dB. For low vehicles the sloping surface degrades the performance

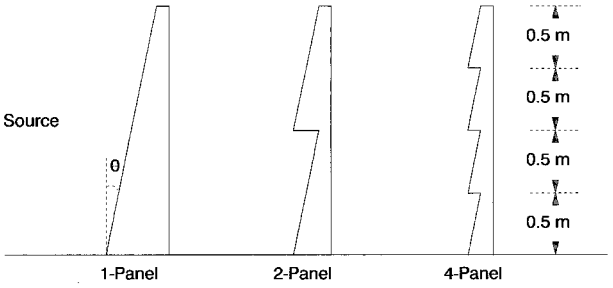


FIG. 7. Cross sections of three types of sloping barriers.

from that of the vertical face. The two-panel design appears to be the most effective, but the detail of the results will be sensitive to the site conditions, the source spectrum, and the source and receiver positions used.

A constant vehicle height of 4.0 m is assumed in Fig. 9 and the mean insertion loss is plotted as a function of the slope angle, θ, for the different designs in Fig. 7. As the value of θ increases from 0° to 20° there is a general increase in mean insertion loss. At angles greater than this the mean insertion loss remains constant; with an increase of approximately 0.5 dB between the four- and two-panel designs and a further 0.5 dB between the two- and one-panel designs. The maximum mean insertion loss achieved by the one-panel design is similar to that observed for a vertical screen with the traffic-facing side absorbing in Fig. 5.

In considering the basic geometry in Fig. 10 it is clear that for a sloping barrier to be efficient the rays from the source which are reflected in the face of the barrier should pass over the top of the vehicle. The limiting rays A and B are shown. For the ray A, via the top edge of the barrier to graze the top of a 4-m-high vehicle the required slope angle is approximately 2°. For ray B, via the lower edge of the barrier to graze the top of the vehicle the slope angle must be approximately 20°. As the slope angle increases from 2° to 20°, an increasing proportion of the reflected rays pass over

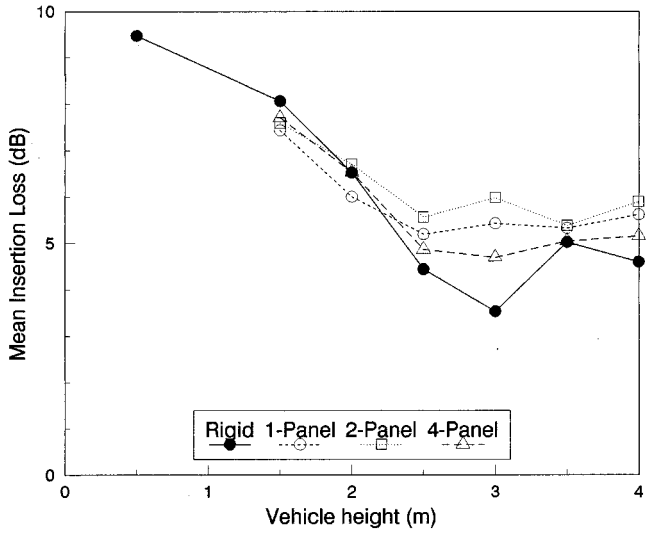


FIG. 8. Mean insertion loss as a function of vehicle height for a 2.0-m rigid barrier with vertical sides and also the forms shown in Fig. 7.

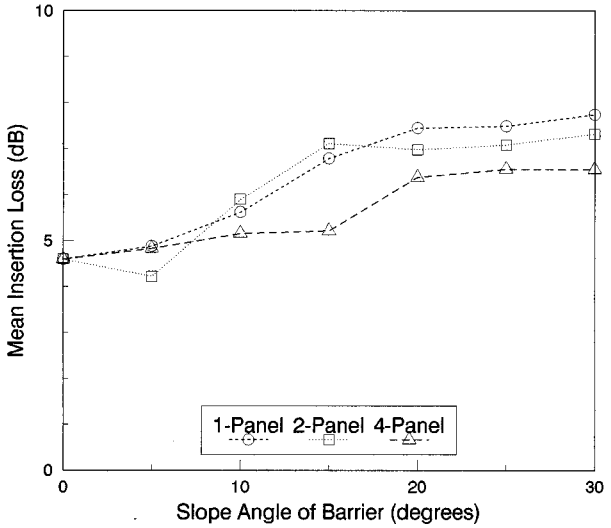


FIG. 9. Mean insertion loss as a function of the slope angle of a 2.0-m rigid barrier with the forms shown in Fig. 7. The vehicle height is constant at 4.0 m.

the vehicle. At 20° all the rays reflected first by the barrier pass over the vehicle. First reflections from the ground surface or vehicle side are not considered here, but this simple argument is sufficient to explain the trends observed in Fig. 9 using the sophisticated wave model.

III. CONCLUSION

A two-dimensional boundary element model has been used in this investigation. This can be interpreted in three dimensions as being equivalent to an infinite line source of sound with an infinite parallel barrier and vehicle of uniform cross section. In terms of road vehicles this is not realistic and the results obtained probably represent a worst case condition. The model is more realistic for the case of railway trains with a barrier near the track and the results could be related to these conditions.

The results show that multiple reflections from the sides of vehicles can significantly degrade the performance of a rigid barrier, particularly when the vehicle is higher than the barrier, by a typical value of 5 dB. The reduction in efficiency did not appear to be sensitive to the separation of the vehicle and barrier up to 22 m. For every configuration considered the insertion loss of the barrier remained positive. This is contrary to the results obtained using a dynamic image source model reported in Ref. 13. Here, when the vehicle

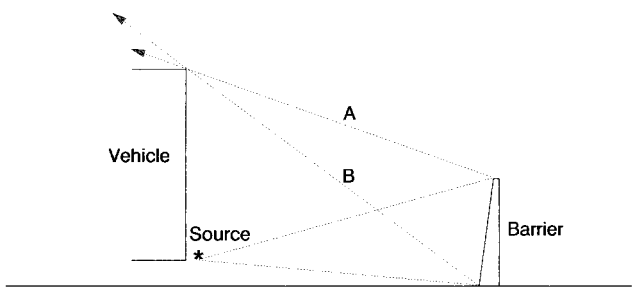


FIG. 10. Limiting reflected rays for a vehicle near a barrier.

is in closest proximity to the receiver the barrier insertion loss reaches a negative value in many cases.

The degradation in performance due to multiple reflections can be reduced by the application of absorbing material on the side of the barrier facing the source. The simulations indicate that the effect depends on the area of the absorbing surface. Provided that the source is in the position used in the model (i.e., well below the upper edge of the barrier) when the whole surface is treated with absorber the effects of multiple reflections can be almost completely removed even when the vehicle side projects above the barrier.

The effects of multiple reflections can be overcome by sloping a rigid barrier. Using a sophisticated wave model it has been confirmed that the minimum slope angle for this to be achieved can be determined approximately by consideration of the limiting ray from the source which is reflected from the lower edge of the barrier grazes the top of the vehicle. When the separation of vehicle and barrier is 5.5 m the minimum angle is approximately 20°. The efficiency of the design increases slightly with the area of the individual sloping panels.

The model assumes that the atmosphere is still and homogeneous. Atmospheric refraction and scattering due to air turbulence could be significant in specific site conditions.

ACKNOWLEDGMENTS

This work was supported by a grant from the Engineering and Physical Sciences Research Council. The contribution of Dr. S. N. Chandler-Wilde in the development of the numerical model is acknowledged.

[1] Calculation of Road Traffic Noise, Department of Transport and Welsh Office, HMSO, London (1988).

[2] J. P. Clairbois, ''Road and rail noise: Corrective devices,'' Seminar on Acoustical Barriers—The Engineered Solution to Road and Rail Noise Pollution, I. Mech. E., London, September 1990.

[3] D. C. Hothersall and S. A. Tomlinson, ''High sided vehicles and road traffic noise barriers,'' Proc. Internoise '95, 397–400 (1995).

[4] D. C. Hothersall and S. A. Tomlinson, ''High-sided vehicles and road traffic noise barriers,'' Acoust. Bull. **21**, 43–46 (1996).

[5] R. Seznec, ''Diffraction of sound around barriers: Use of boundary elements technique,'' J. Sound Vib. **73**, 195–209 (1982).

[6] D. C. Hothersall, S. N. Chandler-Wilde, and M. N. Hajmirzae, ''Efficiency of single noise barriers,'' J. Sound Vib. **146**, 303–332 (1991).

[7] S. N. Chandler-Wilde, D. C. Hothersall, D. H. Crombie, and A. T. Peplow, ''Efficiency of an acoustic screen in the presence of an absorbing boundary,'' Publications du Laboratoire de Mechanique et d'Acoustique, CNRS Marseille, No. 126, 73–90 (1991).

[8] S. N. Chandler-Wilde and D. C. Hothersall, ''Efficient calculation of the Green function for acoustic propagation above a homogeneous impedance plane,'' J. Sound Vib. **180**, 705–724 (1995).

[9] M. E. Delany and E. N. Bazley, ''Acoustical properties of fibrous absorbing materials,'' Appl. Acoust. **3**, 105–116 (1969).

[10] G. R. Watts, D. H. Crombie, and D. C. Hothersall, ''Acoustic performance of new designs of traffic noise barriers–Full scale tests,'' J. Sound Vib. **177**, 289–305 (1994).

[11] C. W. Menge, ''Sloped barriers for highway noise control,'' Proc. Internoise '78, 509–512 (1978).

[12] S. Slutsky and H. L. Bertoni, ''Analysis and programs for assessment of absorptive and tilted parallel noise barriers,'' in Research on Noise and Environmental Issues, Transporation Research Record 1176, Transportation Research Board, National Research Council, Washington, DC, 13–22 (1988).

[13] J. P. Clairbois, P. Houtave, and V. Trefois, ''How multiple reflections can cause a degradation of performance of noise barriers,'' Proc. Internoise '96, 763–768 (1996).

Study of the acoustic feedback problem of active noise control by using the l_1 and l_2 vector space optimization approaches

Mingsian R. Bai and Tianyau Wu
Department of Mechanical Engineering, Chiao-Tung University, 1001 Ta-Hsueh Road, Hsin-Chu, Taiwan, Republic of China

(Received 15 July 1996; revised 10 March 1997; accepted 8 April 1997)

This study attempts to explore the acoustic feedback problem that is frequently encountered in feedforward active noise control (ANC) structure from the standpoint of control theories. The analysis is carried out on the basis of the Youla's parametrization and the l_1-norm and l_2-norm vector space optimization. The ANC problem is formulated as a model matching procedure and is solved via linear programming in the dual vector space. These methods alleviate the problems caused by the nonminimum phase (NMP) zeros of the cancellation path. The ANC algorithms are implemented by using a digital signal processor. Various types of noises were chosen for validating the developed algorithms. The proposed methods are also compared with the well-known filtered-u least-mean square (FULMS) method. The experimental results show that the acoustic feedback problem significantly degrades the performance and stability of the ANC system regardless which method is used. Insights into the difficulties due to acoustic feedback are addressed along the line of control theories. © *1997 Acoustical Society of America.* [S0001-4966(97)03008-7]

PACS numbers: 43.50.Ki [GAD]

INTRODUCTION

The active noise control (ANC) technique has attracted much academic as well as industrial attention since it provides numerous advantages over traditional passive methods in attenuating low-frequency noises.[1,2] However, the optimistic view somewhat masks the complexity of many theoretical and technical problems that remain to be solved prior to full commercialization of the technique. Among the difficulties, the acoustic feedback problem of the feedforward structure has been the plaguing issue that seriously impedes successful application of the ANC technique. This problem usually arises in a feedforward ANC structure that is generally the only practical approach for suppressing broadband random noises, where in many practical situations only the upstream acoustical reference is available. In this configuration, a positive feedback loop will exist between the canceling loudspeaker and the feedforward microphone, which tends to destabilize the ANC system. Consider the ANC system for a duct shown in Fig. 1. Suppose that only the noise below the cutoff frequency of the duct is of interest so that the control problem can be treated as a one-dimensional problem. The secondary loudspeaker generates a canceling sound to interact destructively with the primary noise field. It is desired that the noise will be attenuated as much as possible at the error microphone position. It is further assumed that only the acoustical reference at the upstream of the canceling source is available, as usually is the case in practical applications. Unfortunately, the canceling loudspeaker on a duct wall will generate plane waves propagating both upstream and downstream. Therefore, the antisound output to the loudspeaker not only cancels noise downstream by minimizing the error signals measured by the error microphone but also radiates upstream to the reference microphone, resulting in a corrupted reference signal. The coupling of the acoustic wave from the canceling loudspeaker to the reference microphone is called *acoustic feedback*.[3] Similar effects take place in the active vibration control system due to feedback from the control actuator to the reference sensor. Because the upstream feedforward microphone will detect both the noise from the primary source and the noise produced by the canceling loudspeaker, the positive acoustic feedback problem can no longer be ignored. It is this positive feedback problem that seriously complicates the active control design.

The fact that acoustic feedback introduces poles to the system and thus a stability problem if the loop gain becomes too large has been illustrated by Eghtesadi and Leventhall.[1,4] Another good review concerning acoustic feedback can be found in Chapter 3 of the book by Kuo and Morgan.[3] The solutions to the problem of acoustic feedback cited in the book included directional microphones and loudspeakers, motional feedback loudspeakers, neutralization filter, dual-microphone reference sensing, the filter-u least-mean square (FULMS) method, nonacoustic sensors, and so forth.

This paper adopts a new approach to explore the acoustic feedback problem in the ANC application domain from the standpoint of control theories. The analysis is carried out on the basis of the Youla's parametrization[5] and the l_1-norm and l_2-norm vector space optimization.[6] The performance of the developed methods is compared with the well-known FULMS algorithm that has long been recognized as one of the elegant solutions to the acoustic feedback problem.[3,7] The simplicity of the proposed methods stems from the fact that the ANC problem is formulated as a model matching procedure and is solved via linear programming (LP) in the dual vector space. These methods optimally tackle the problem of unstable inverse resulting from the inherent nonminimum phase (NMP) property of the cancellation path (which frequently occurs to structural systems or acoustical systems).

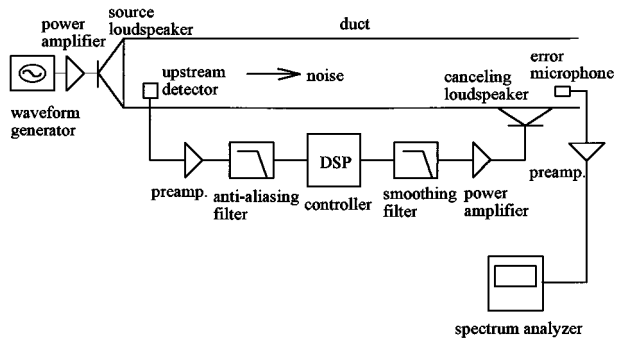


FIG. 1. Schematic diagram of the feedforward ANC structure for a duct in the presence of acoustic feedback.

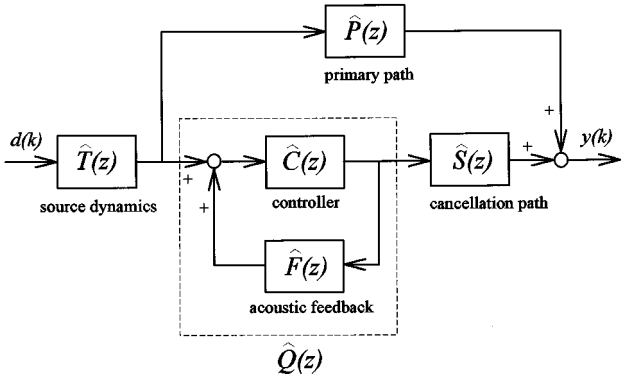


FIG. 2. System block diagram of the feedforward ANC structure in the presence of acoustical feedback. $\hat{P}(z)$, $\hat{S}(z)$, $\hat{F}(z)$, and $\hat{C}(z)$ are the discrete time transfer functions of the primary path, the cancellation path, the acoustic feedback path, and the controller, respectively.

The l_1 and l_2 ANC algorithms are implemented by using a floating-point digital signal processor. Various types of noises including white noise, engine exhaust noise, and sweep sine were chosen for validating the developed algorithms. It will be seen from the experimental results that the acoustic feedback problem indeed significantly degrades the performance and stability of the ANC system regardless of which ANC method is used. Additional insights into the difficulties due to acoustic feedback are also addressed in this paper within a general framework of control theories.

I. THE l_1 AND l_2 ANC ALGORITHMS

The l_1 control theory emerged in the 1980s. Dahleh and Pearson proposed an l_1 compensator to track persistent bounded signals.[8] The same authors also applied the l_1 control theory to reject bounded persistent disturbances by solving a semi-infinite LP problem.[9] Deodhare and Vidyasagar performed an in-depth analysis of the condition of when a stabilizing controller is l_1 optimal.[10] An excellent review of the l_1 control theory can be found in the text by Dahleh and Diaz-Bobillo.[6] In the sequel, the l_1 control theory is only presented in the context of the ANC problem.

As mentioned previously, the acoustic feedback problem is very detrimental in terms of performance and stability of the ANC system. In what follows, the analysis of acoustic feedback problem in the ANC application domain is carried out from the standpoint of control theories. To begin with, the ANC problem is formulated as a model matching problem,[11] as shown in the block diagram of Fig. 2. The sampled signal $d(k)$ is the input disturbance noise and $y(k)$ is the residual field detected at the error microphone position. The primary path, $\hat{P}(z)$, is the transfer function formed by the primary acoustic path, the power amplifier, the upstream sensor, and the error microphone, where z denotes the z transform variable. $\hat{S}(z)$ is the transfer function of the cancellation path formed by the canceling loudspeaker, the power amplifier, the secondary acoustic path, and the error microphone. $\hat{F}(z)$ is the transfer function of the acoustic feedback path formed by the canceling loudspeaker, the power amplifier, the acoustic path, and the upstream sensor. The transfer function $\hat{C}(z)$ represents the controller to be designed. In particular, $\hat{T}(z)$ is a transfer function describing the source dynamics formed by the primary source, a power amplifier, an acoustical path, and the upstream microphone. Hence, the feedforward ANC problem in the presence of acoustical feedback amounts to finding the controller $\hat{C}(z)$ such that the residual field $y(k)$ can be minimized. This is a generic model matching problem. That is,

$$\min_{C(z)} \left\| T(z)\left[P(z)+S(z)\,\frac{C(z)}{1-F(z)C(z)}\right]\right\|, \tag{1}$$

where $\|(.)\|$ is the notation of function norm. Note that the transfer function, $\hat{T}(z)$, can be regarded as a weighting function for the terms inside the parenthesis. In practice, $\hat{T}(z)$ may not be measurable so that either analytical modeling or heuristic guess is needed.

By Youla's parametrization, the controller $\hat{C}(z)$ that guarantees the internal stability of the closed loop system can be expressed as[12]

$$\hat{C}(z)=\frac{\hat{M}(z)\hat{Q}(z)+\hat{Y}(z)}{\hat{N}(z)\hat{Q}(z)+\hat{X}(z)}, \tag{2}$$

where $\hat{Q}(z)$ is a stable, proper, and real-rational function [denoted as $\hat{Q}(z)\in RH^{\infty}$],[12] $\hat{F}(z)=\hat{N}(z)/\hat{M}(z)$ is the coprime factorization of $\hat{F}(z)$, and $\hat{X}(z)\in RH^{\infty}$ and $\hat{Y}(z)\in RH^{\infty}$ satisfy the Bezout identity[12]

$$\hat{M}(z)\hat{X}(z)-\hat{N}(z)\hat{Y}(z)=1. \tag{3}$$

If the plant is proper and stable, as it is in our case, one may simply let[12] $\hat{M}(z)=1$, $\hat{N}(z)=\hat{F}(z)$, $\hat{X}(z)=1$, and $\hat{Y}(z)=0$. Then the controller in Eq. (2) turns out to be

$$\hat{C}(z)=\frac{\hat{Q}(z)}{\hat{F}(z)\hat{Q}(z)+1}. \tag{4}$$

Substituting Eq. (4) into Eq. (1) leads to

$$\min_{Q(z)\in RH^{\infty}} \|T(z)[P(z)+S(z)Q(z)]\|. \tag{5}$$

It is noteworthy that the model matching problem in Eq. (5) is exactly of the same form as that of the purely feedforward ANC system (in the absence of acoustic feedback) by replacing the designed controller $\hat{C}$ by a parameter $\hat{Q}$ that is indicated by the dashed line in Fig. 2. Since $\hat{S}(z)$ is usually

NMP, it is desirable to develop an algorithm that would effectively match the primary path and the cancellation path with respect to some optimal criterion, especially when the NMP behavior is present. Elegant solutions to this problem are available in the robust control theory. Two model matching algorithms capable of dealing with NMP problems in ANC applications are presented in this paper. The optimal solution of Eq. (5), $\hat{Q}_{\text{opt}}(z)$, can be found by either l_1 or l_2 optimization procedure.

$$\min_{Q\in l_1}\|P+S*Q\|_1=\min_{H\in l_1}\|G+H\|_1$$

or (6)

$$\min_{Q\in l_2}\|P+S*Q\|_2=\min_{H\in l_2}\|G+H\|_2,$$

where P and S are pulse response sequences of the primary path and the cancellation path, respectively, ''*'' represents convolution, $G=P$, $H=S*Q$, ''$\|\cdot\|_{1,2}$'' denotes the 1-norm or 2-norm, and l_1 denotes the normed linear space of all bounded sequences with bounded 1-norm.[13]

The following derivation contains a fair amount of mathematical definitions and results. However, the l_1 theories are quite standard in functional analysis[14] and vector space optimization,[13] and we present only the key ones needed in the development of the l_1-ANC algorithm. The rest are mentioned without proof.

The minimum distance problem in the primal space posed in Eqs. (6) can be recast into a maximum problem in its dual space. Let X be a normed linear space; f is said to be a bounded linear functional on X if f is a continuous linear operator from X to $\Re$. A dual space of X, denoted by X^*, is the collection of all bounded linear functionals on X, equipped with the natural induced norm. Whenever f is represented by some element $x^*\in X^*$, we use the notation $\langle x,x^*\rangle$ to represent the value $f(x)$. Then, a vector is said to be *aligned* with its normed dual if the following relation is satisfied[13]

$$\langle x,x^*\rangle=\|x^*\|\|x\|. \tag{7}$$

Given a sequence $h=[h(1)\ h(2)\ h(3)\ \cdots]'$ in l_1, where the prime denotes matrix transpose, we also need the definition of λ transform of h to facilitate the following derivations of l_1 algorithm.

$$\hat{H}(\lambda)=\sum_{k=0}^{\infty}h(k)\lambda^k. \tag{8}$$

Unlike the Z transform, it should be noted that the λ transform is defined in terms of an infinite series of positive powers. It is obvious that $\hat{H}(\lambda)$ is BIBO (bounded-input–bounded-output, i.e., both the input and the output are l_∞ sequences) stable if and only if h is an l_1 sequence.[8] In addition, the space c_0 is defined as the space of all infinite sequences in l_∞ of real numbers converging to zero.

With the aforementioned definitions, we are in a position to find the optimal solution of the model matching problem in Eqs. (6). Assume that $\hat{S}(\lambda)$ has no zeros on the unit circle. Then the so-called *interpolation constraints* must be satisfied.[8] Let $\hat{S}(\lambda)$ have N NMP zeros denoted by a_i, $i=1,2,\cdots,N$, and $|a_i|<1$, each of multiplicity γ_i, then

$$\hat{H}^{(k)}(a_i)=0,\quad i=1,2,\ldots,N;\quad k=0,1,\ldots,\gamma_i-1, \tag{9}$$

where the superscript (k) denotes the kth derivative of a complex function. The key of the solution to the optimization problem in Eqs. (6) is based on the following *duality theorem*.[15]

Theorem: Given the system of m consistent linear equations in n unknowns $Ax=y$, then

$$\min_{Ax=y}\|x\|_p=\max_{\|A'u\|_q\leqslant 1}(y'u), \tag{10}$$

where $1/p+1/q=1$, $1<p,\ q<\infty$. Furthermore, optimal x and $A'u$ are aligned. According to the theorem, the primal problem (which is a minimization problem) can be converted to a maximization problem in its dual space. By the duality theorem, it can be shown that the l_1 optimization problem in Eq. (6) becomes[6]

$$d_{\text{opt}}=\|\phi_{\text{opt}}\|_1=\min_{V_\infty'\phi=\Gamma}\|G+H\|_1=\max_{\|V_\infty x\|_\infty\leqslant 1}(\Gamma'x), \tag{11}$$

where d_{opt} is the minimal distance, $\phi=G+H=[\alpha_1\ \alpha_2\ \alpha_3\cdots]'$ is the error vector, ϕ_{opt} is the optimal ϕ,

$$\Gamma=[\hat{G}(a_1)\hat{G}^{(1)}(a_1)\cdots\hat{G}^{(\gamma_1)}(a_1)\hat{G}(a_2)\times\hat{G}^{(1)}(a_2)\cdots\hat{G}^{(\gamma_2)}(a_2)\cdots\hat{G}^{(\gamma_N)}(a_N)]',\ x\in\Re^m,$$

and

$$V_\infty=\begin{bmatrix}\text{Re}(a_1^0) & \ldots & \text{Re}(a_N^0) & \text{Im}(a_1^0) & \ldots & \text{Im}(a_N^0)\\ \text{Re}(a_1^1) & \ldots & \text{Re}(a_N^1) & \text{Im}(a_1^1) & \ldots & \text{Im}(a_N^1)\\ \vdots & & \vdots & \vdots & & \vdots\end{bmatrix}.$$

In the matrix V_∞, Re.$(\cdot)$ and Im$(\cdot)$ denote the real part and the imaginary part, respectively. Note also that the interpolation constraints in Eq. (9) has been rewritten as $V_\infty'\phi=\Gamma$. Hence the infinite-dimensional optimization problem has become a finite-dimensional one. The only pitfall is that the LP problem in Eq. (11) has infinitely many constraints. However, because $|a_i|<1$, it can be shown that only a finite number of these constraints are really necessary.[6] More precisely, we can compute an integer, L, such that any α_i satisfies the first L constraints also satisfies the remaining ones. Define a real Vandermonde matrix of order L

$$V_L=\begin{bmatrix}\text{Re}(a_1^0) & \ldots & \text{Re}(a_N^0) & \text{Im}(a_1^0) & \ldots & \text{Im}(a_N^0)\\ \text{Re}(a_1^1) & \ldots & \text{Re}(a_N^1) & \text{Im}(a_1^1) & \ldots & \text{Im}(a_N^1)\\ \vdots & & \vdots & \vdots & & \vdots\\ \text{Re}(a_1^L) & \ldots & \text{Re}(a_N^L) & \text{Im}(a_1^L) & \ldots & \text{Im}(a_N^L)\end{bmatrix}. \tag{12}$$

Hence, the matrix V_∞ in Eq. (11) can everywhere be replaced by V_L. If a_i is real, then the column of zeros corresponding to the imaginary part is dropped. If a_i is complex, either a_i or its complex conjugate $\overline{a_i}$ is considered. Note that the matrix V_L has full column rank for all $L\geqslant 2N$. An upper bound on

the number of the constraints, $L(\geqslant 2N)$, can be computed *a priori*:[6]

$$\max_i |a_i|^L (L|V_L^{-l}|_1) < 1, \quad i = 1,2,\ldots,N, \tag{13}$$

where V_L^{-l} is the left inverse of V_L.

The beauty of the above l_1 approach lies in the fact the infinite-dimensional model matching problem in Eq. (11) has been converted to a standard LP problem described by

$$d_{\text{opt}} = \max_{\|V_L x\| \leqslant 1} (\Gamma' x), \tag{14}$$

which can be conveniently solved in the dual space by using an LP routine. Then the corresponding optimal solution is recovered in the primal space via the alignment conditions. Define

$$\phi_{L,\text{opt}} = [\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_L]' \in l_1$$

and

$$b = [\beta_1 \quad \beta_2 \quad \cdots \quad \beta_L]' = V_L x_{\text{opt}} \in \mathscr{C}_0.$$

The l_1 version of the alignment conditions are stated as:[6,15] (a) $\alpha_i = 0$ if $|\beta_i| < 1$; (b) $\alpha_i \beta_i \geqslant 0$ if $|\beta_i| = 1$. However, only part (a) of the alignment conditions in conjunction with the interpolation constraints, $V_L' \phi = \Gamma$, are needed for determining the components of $\phi_{L,\text{opt}}$. After $\phi_{L,\text{opt}}$ has been solved, the finite-impulse response function $\hat{\phi}_{L,\text{opt}}(z)$ can be expressed (by using the fact $\lambda = z^{-1}$) as

$$\hat{\phi}_{L,\text{opt}}(z) = \alpha_1 + \alpha_2 z^{-1} + \cdots + \alpha_L z^{-(L-1)}. \tag{15}$$

By the same token, the feedforward ANC problem can be formulated as an l_2 model matching problem. The minimization problem in the primal space can be converted to a maximization problem in the dual space by the duality theorem:

$$\begin{aligned} d_{\text{opt}} &= \|\phi_{L,\text{opt}}\|_2 = \min_{\mathbf{V}_{L^*}' \Phi = \Gamma} \|G - H\|_2 = \max_{\|\mathbf{V}_{L^*} \mathbf{x}\|_2 \leqslant 1} [\Gamma' \mathbf{x}], \\ \Gamma &= [\hat{G}(a_1)\hat{G}^{(1)}(a_1)\cdots\hat{G}^{(\gamma_1)}(a_1)\hat{G}(a_2) \\ &\quad \times \hat{G}^{(1)}(a_2)\cdots\hat{G}^{(\gamma_2)}(a_2)\cdots\hat{G}^{(\gamma_N)}(a_N)]', \end{aligned} \tag{16}$$

where the notations are the same as those defined in Eq. (11). It can be shown that the alignment condition leads directly to the optimal solution[5,14]

$$\phi_{L,\text{opt}} = V_L (V_L' V_L)^{-1} \Gamma \tag{17}$$

with the corresponding minimum distance

$$d_{\text{opt}} = [\Gamma' (V_L' V_L)^{-1} \Gamma]^{1/2}. \tag{18}$$

Note that the inverse of $(V_L' V_L)$ is guaranteed since V_L is of full column rank. The only difference between the l_2 optimization and the l_1 optimization is that the former does not require explicit solution of the dual problem.

Next, both the l_1 and l_2 optimal parameters, $\hat{Q}_{\text{opt}}(z)$, can be obtained by backsubstituting $\hat{\phi}_{L,\text{opt}}(z)$ into Eq. (11) or Eq. (16) as follows:

$$\hat{Q}_{\text{opt}}(z) = [\hat{\phi}_{L,\text{opt}}(z) - \hat{P}(z)]/\hat{S}(z). \tag{19}$$

It should be noted that, because the interpolation constraints have been satisfied, the NMP zeros of $\hat{S}(z)$ will always be canceled by $\hat{\phi}_{L,\text{opt}}(z) - \hat{P}(z)$ to ensure the stability of the parameter $\hat{Q}_{\text{opt}}(z)$. After $\hat{Q}_{\text{opt}}(z)$ is obtained, the optimal controller $\hat{C}_{\text{opt}}(z)$ can be recovered by the following formula:

$$\hat{C}_{\text{opt}}(z) = \frac{\hat{Q}_{\text{opt}}(z)}{\hat{F}(z)\hat{Q}_{\text{opt}}(z) + 1}. \tag{20}$$

A note regarding the implementation of the above l_1 controller and l_2 controller is in order. In some cases, it may occur that these methods result in controllers with excessive gains in the high-frequency range. This is due to the fact that the control algorithms attempt to compensate for the NMP zeros in high frequencies that may be only artificially created from discretization.[16] To alleviate the problem, remedies may be sought by incorporating a small constant ϵ into $\hat{S}(z)$, i.e.,

$$S(z) \rightarrow [S(z) + \epsilon] \tag{21}$$

so that the plant is not excessively rolled off.

In some cases, the resulting controller $\hat{C}_{\text{opt}}(z)$ can be unstable, even though the parameter $\hat{Q}_{\text{opt}}(z)$ is guaranteed to be proper and stable. This fact has been illustrated in an acoustic analysis in frequency domain of a simplified infinite-length duct model detailed in the paper by Eghtesadi and Leventhall.[4] An alternative argument from the viewpoint of control theory will lead to similar observation as follows. Let the transfer function between the disturbance $d(k)$ and the residual output $y(k)$ be $\hat{P}_1(z)$. Recall that $d(k)$ is assumed to be inaccessible and only $\hat{P}(z) = \hat{P}_1(z)\hat{T}^{-1}(z)$ is subject to some sort of experimental system identification procedure. However, it happens more often than never that the system, $\hat{T}(z)$, is NMP such that $\hat{T}(z)$ has no stable inverse. Direct identification of the plant $\hat{P}(z)$ generally results in unstable or nearly unstable systems, even though both $\hat{P}_1(z)$ and $\hat{T}(z)$ are stable. Hence, it is likely that the identified $\hat{P}_1(z)$ involves large modeling error and the best one can get is only an approximation of the inherently unstable $\hat{T}^{-1}(z)$. This approximation further degrades the performance of active control in comparison with the purely feedforward structure. Nevertheless, this modeling difficulty can be alleviated somewhat by appropriate introduction of passive damping. The importance of passive damping treatment that has often been overlooked in active design lies in not only high-frequency attenuation but also the robustness of active control with respect to plant uncertainties.[17] Another benefit of using passive damping is that lower order of plant models can usually be obtained than the lightly damped plants so that numerical error is reduced.

II. EXPERIMENTAL RESULTS

Experiments were conducted to validate the adaptive l_1 and l_2 ANC algorithms. A rectangular duct of length 1 m and cross section $0.25\ \text{m} \times 0.25\ \text{m}$ was chosen for this test. The duct is lined with fiberglass sound-absorbing material to create sufficient acoustical damping that is crucial to the ro-

TABLE I. The poles and zeros of the primary path $\hat{P}(z)$, the cancellation path $\hat{S}(z)$, and the acoustic feedback path $\hat{F}(z)$.

$\hat{P}(z)$		$\hat{S}(z)$		$\hat{F}(z)$	
Zeros	Poles	Zeros	Poles	Zeros	Poles
9.9496	−0.9028±	0	0.9603±	0	−0.7521±
−1.7015	0.1998*i*	0	0.1314*i*	0	0.4749*i*
−0.9043	−0.9801	0	0.7822±	0	−0.8119±
−0.9067±	−0.8536	−13.9155	0.4891*i*	2.0484±	0.2015*i*
0.1871*i*	−0.7881±	1.0253	0.4660±	1.8137*i*	−0.4763±
−0.7956±	0.3770*i*	0.9604	0.7756*i*	−1.0837±	0.7045*i*
0.3710*i*	−0.7128±	0.6763±	0.5821±	2.0421*i*	−0.3262±
−0.7101±	0.5819*i*	0.6445*i*	0.5469*i*	1.0389	0.7085*i*
0.5863*i*	−0.5956±	0.3838±	0.1132±	0.9233	−0.0983±
−0.5852±	0.6923*i*	0.7692*i*	0.8661*i*	0.6327±	0.8864*i*
0.7100*i*	−0.4439±	0.1332±	−0.9756	0.6693*i*	0.3223±
−0.4414±	0.7825*i*	0.8898*i*	−0.2417±	0.3143±	0.8489*i*
0.7944*i*	−0.3999±	−0.2228±	0.8059*i*	0.7870*i*	0.2237±
−0.3101±	0.7295*i*	0.7830*i*	−0.7367±	−1.0706	0.7123*i*
0.8503*i*	−0.2992±	−0.7419±	0.3400*i*	−0.8401±	0.6590±
−0.1585±	0.8666*i*	0.2637*i*	−0.5724±	0.3254*i*	0.6588*i*
0.9111*i*	−0.1497±	0.5260±	0.5447*i*	−0.0532±	−0.7800±
0.0666±	0.9207*i*	0.4643*i*	−0.3582±	0.7865*i*	0.4911*i*
0.9405*i*	0.0704±	0.4790	0.6160*i*	−0.4994±	0.8501±
0.2757±	0.9221*i*	−0.5067	0.6335	0.4428*i*	0.2237*i*
0.8730*i*	0.9975			−0.2340±	0.9579±
0.5087±	0.8880±			0.6325*i*	0.1476*i*
0.8030*i*	0.2575*i*				−0.7250
0.6232±	0.8813±				
0.7250*i*	0.3695*i*				
1.0075	0.7996±				
0.8902±	0.4107*i*				
0.2599*i*					
0.7559±	0.7463±				
0.5024*i*	0.5074*i*				
0.8051±	0.6792±				
0.4125*i*	0.6363*i*				
0.5630±	0.2653±				
0.1642*i*	0.8560*i*				
	0.5036±				
	0.7872*i*				
	0.4152±				
	0.6802*i*				
delay=3,	gain=−0.0247	delay=0,	gain=0.0720	delay=1,	gain=0.0064

bustness of the ANC system, as has been mentioned previously. The corresponding cutoff frequency of the duct is 690 Hz and thus only plane waves are of interest. The controller was implemented by using a floating-point TMS320C31 digital signal processor. The sampling frequency was chosen to be 2 kHz. Fourth-order Butterworth filters with cutoff frequency 600 Hz were used as the anti-aliasing filter and the smoothing filter. The experimental setup including the duct, the upstream sensor, the error microphone, the actuator, the controller, and signal conditioning circuits has been shown in Fig. 1. The ARX system identification procedure[18] was used for establishing the mathematical models of the primary path, the cancellation path, and acoustic feedback path whose poles and zeros are listed in Table I. The frequency response functions of the primary path, the cancellation path, and the acoustic feedback path are shown in Figs. 3–5. The frequency response function of the chosen weighting function in Eq. (5) is shown in Fig. 6. The frequency response

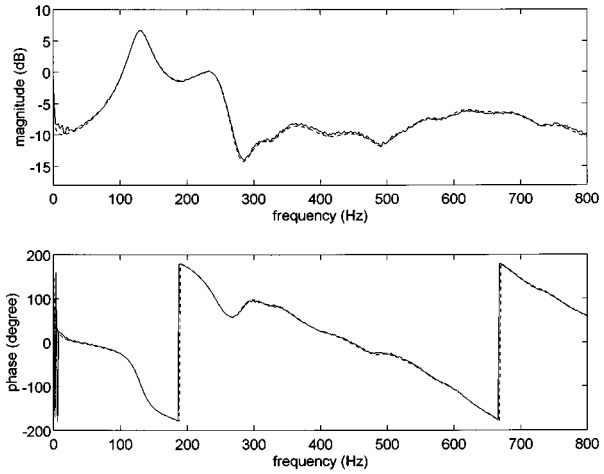


FIG. 3. The frequency response function of the primary path $\hat{P}(z)$. Measured data: ———; model data: -----.

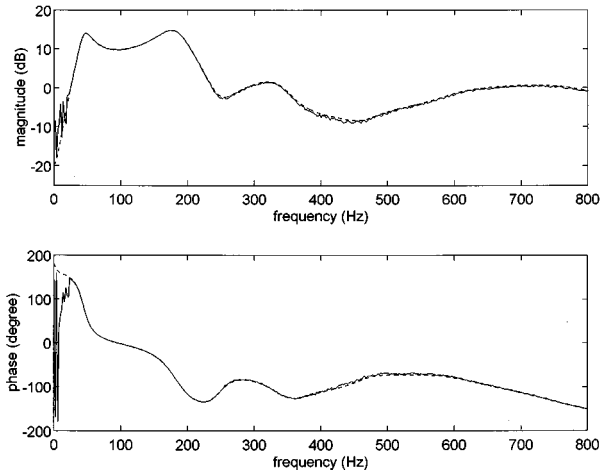


FIG. 4. The frequency response function of the cancellation path $\hat{S}(z)$. Measured data: ———; model data: -----.

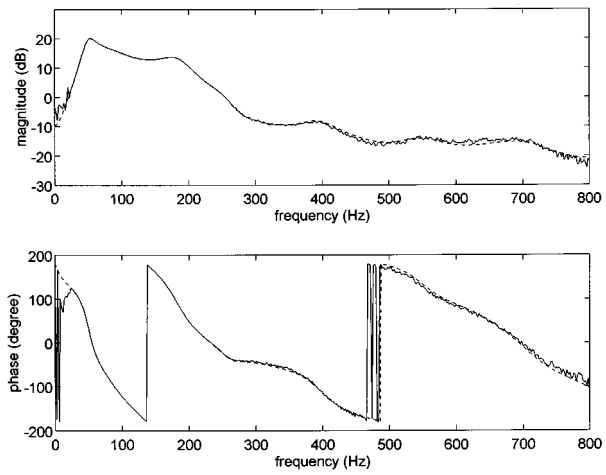


FIG. 5. The frequency response function of the acoustic feedback path plant $\hat{F}(z)$. Measured data: ———; model data: -----.

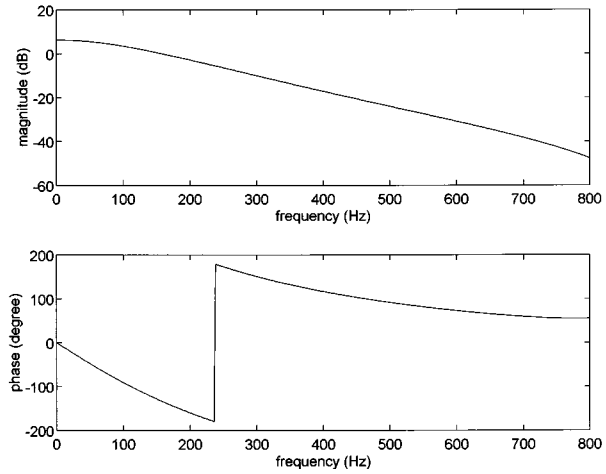


FIG. 6. The frequency response function of the chosen weighting function $\hat{T}(z)$ in Eq. (5).

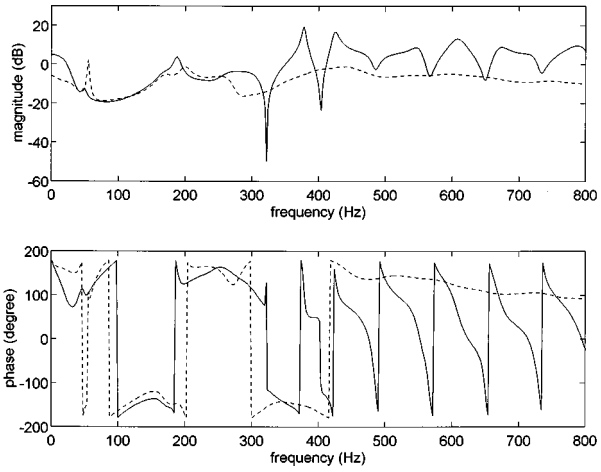


FIG. 7. The frequency response functions of the l_1 and l_2 controllers in the experimental cases of Table II. l_1 controller: ———; l_2 controller: -----.

functions of the optimal controllers designed by the aforementioned l_1 and l_2 procedure are then shown in Fig. 7.

To compare the performance of the ANC techniques, three types of noises including a white noise, a sweep sine, an engine exhaust noise were employed as the primary noises. The experimental cases and the associated noise types and control algorithms are summarized in Table II.

Case 1 is a reference case of purely feedforward l_1 control, where the source voltage signal is used as the feedforward reference so that acoustic feedback can be neglected. It can be seen from the result of Fig. 8 that significant attenuation (approximately 5–20 dB) is obtained in the frequency range 60–330 Hz. In cases 2–4, the signal detected by an upstream microphone is used as the feedforward reference so that the effect of acoustic feedback is present. In these cases, the widely used FULMS algorithm, the l_1 algorithm, and the l_2 algorithm are employed for attenuating the white noise. The power spectra of the residual fields at the error microphone position with and without active control are shown in Figs. 9, 10, and 11 for cases 2, 3, and 4, respectively. Attenuation (approximately 3–20 dB) was obtained from the FULMS method in the frequency range 50–200 Hz. The result was obtained 1 min after the active control was activated. However, in Fig. 9, a strong peak was found at 235 Hz in the spectrum with the FULMS algorithm. On the other hand, attenuation (approximately 2–30 dB) was obtained from the l_1 and l_2 method in the frequency range 50–220 Hz. Total attenuation within the dominant band 0–400 Hz is found to be 1.96 dB, 2.31 dB, and 2.93 dB for the FULMS method, the l_1 method, and the l_2 method, respectively. Although the total attenuation appears to be comparable, these three control algorithms give different emphasis for different frequency ranges. The FULMS method produced most attenuation (approximately 20 dB) around 100 Hz, whereas the l_1 and l_2 methods produced most attenuation (approximately 30 dB) around 200 Hz. The FULMS method is prone to have a stability problem, as compared with the l_1 method, and the

TABLE II. Experimental case design.

Case	Noise type	Algorithm	Tap/step[a]	Figure
1[b]	white noise	l_1	···	8
2	white noise	FULMS	30/0.0001	9
3	white noise	l_1	···	10
4	white noise	l_2	···	11
5	sweep sine	FULMS	30/0.0001	12
6	sweep sine	l_1	···	13
7	engine exhaust noise	FULMS	30/0.0001	14
8	engine exhaust noise	l_1	···	15

[a]''Tap/step'' denotes the tap length and the step size used in the FULMS algorithm.
[b]Case 1 is a reference case of purely feedforward control.

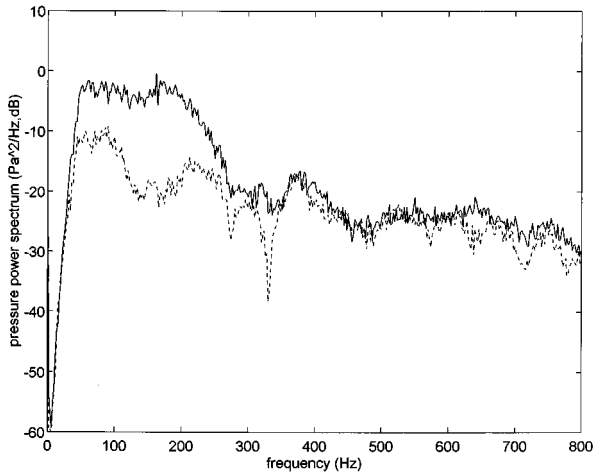


FIG. 8. The residual sound-pressure spectra of case 1 for the white noise in the absence of acoustic feedback before and after ANC is activated by using the purely feedforward l_1 controller. Control off: ———; control on: -----.

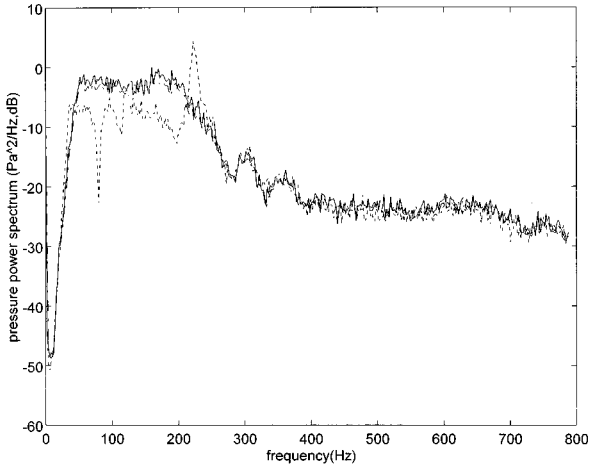


FIG. 9. The residual sound-pressure spectra of case 2 for the white noise in the presence of acoustic feedback before and after ANC is activated by using the FULMS and FXLMS controllers. The step sizes in FULMS and FXLMS are both 0.0001. The filter length in FULMS and FXLMS are 30 and 128, respectively. Control off: ———; FULMS: ---; FXLMS: –·–·.

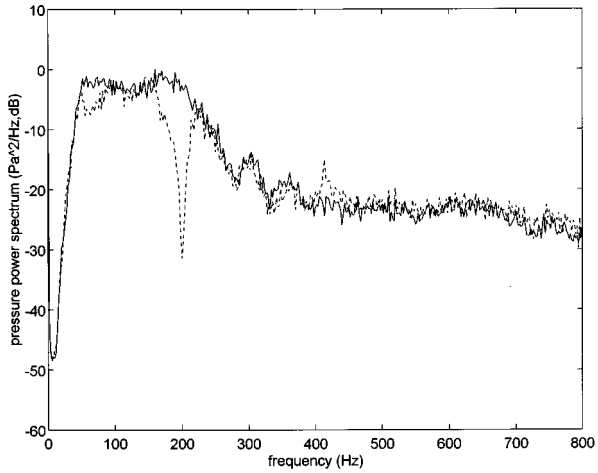


FIG. 10. The residual sound-pressure spectra of case 3 for the white noise in the presence of acoustic feedback before and after ANC is activated by using the l_1 controller. Control off: ———; control on: -----.

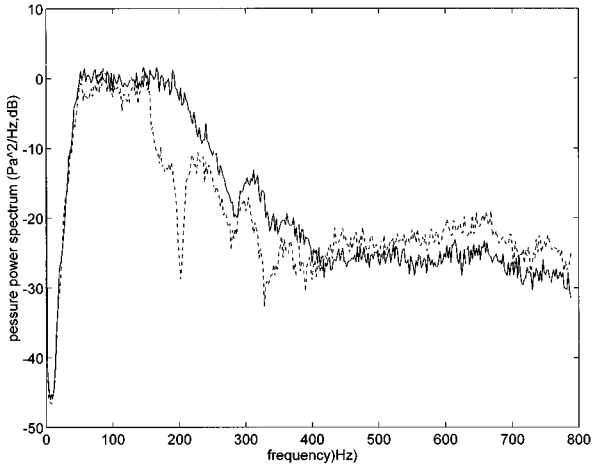


FIG. 11. The residual sound-pressure spectra of case 4 for the white noise in the presence of acoustic feedback before and after ANC is activated by using the l_2 controller. Control off: ———; control on: -----.

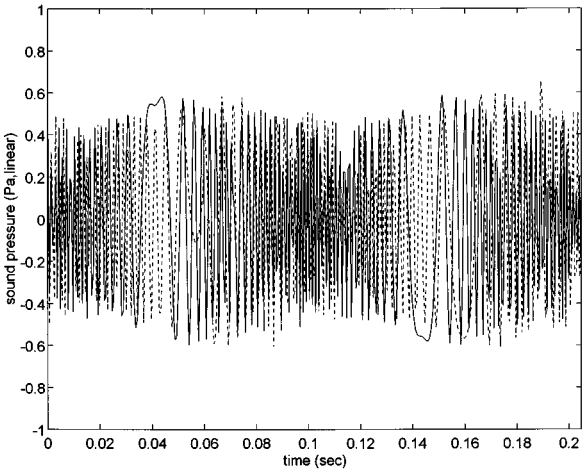


FIG. 12. The time-domain response of case 5 for the sweep sine in the presence of acoustic feedback before and after ANC is activated by using the FULMS controller. The step size in FULMS is 0.0001. The filter length is 30. Control off: ———; control on: -----.

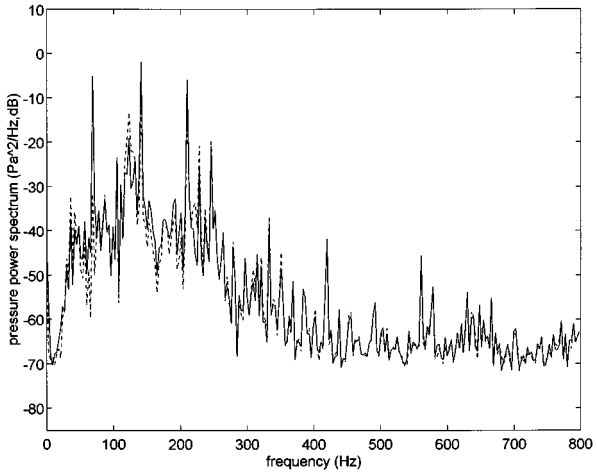


FIG. 14. The residual sound-pressure spectra of case 7 for the engine exhaust noise in the presence of acoustic feedback before and after ANC is activated by using the FULMS controller. The step size in FULMS is 0.0001. The filter length is 30. Control off: ———; control on: -----.

control drifts away slowly so that the system may become unstable. This is probably because the performance index of the FULMS algorithm is not quadratic and the computation is easily trapped in local minima. In any rate, the detrimental effect of acoustic feedback is evidenced from these results. Regardless of which method is used, the performance of noise attenuation is significantly degraded in comparison with the purely feedforward control in case 1. For the sake of comparison, the filtered-x least-mean square (FXLMS) method that is known to be effective for the purely ANC problems without acoustic feedback is also employed to tackle the same case. The result of Fig. 9 shows virtually no cancellation for the method. This is not totally unexpected since the optimal solution of the adaptive filter is generally an IIR function with poles and zeros when acoustic feedback is present.[3] This rational function can only be approximated by an FIR function of infinitely long order. In practice, finite order of the FIR filter has to be used in conjunction with a very small step size for stability reasons. This results in extremely slow convergence in real applications. In the following cases, the results of FXLMS is thus omitted for brevity.

In contrast to cases 2–4 that deal mainly with stationary noises, a sweeping sinusoidal noise is chosen in cases 5 and 6 to compare the effectiveness of the FULMS method and the l_1 method in suppressing a nonstationary noise. The time-domain results of Figs. 12 and 13 show that the l_1 method gives better performance than the FULMS method. This result is expected since the l_1 controller is essentially a fixed controller that does not require adaptation time like the FULMS controller.

In cases 7 and 8, a practical noise, the engine exhaust noise, is employed as the primary noise for comparing the FULMS method and the l_1 method. The power spectra of the residual fields at the error microphone position with and without active control are shown in Figs. 14 and 15. The result of the FULMS method was obtained 1 min after the active control was activated. Total attenuation within the dominant band 0–500 Hz is found to be 6.65 dB and 2.73 dB

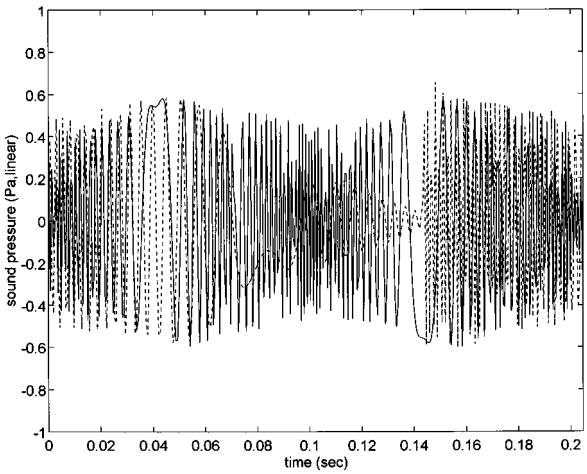


FIG. 13. The time-domain response of case 6 for the sweep sine in the presence of acoustic feedback before and after ANC is activated by using the l_1 controller. Control off: ———; control on: -----.

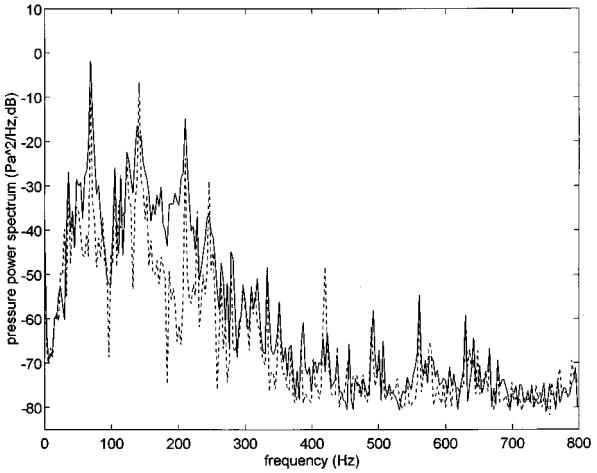


FIG. 15. The residual sound-pressure spectra of case 8 for the engine exhaust noise in the presence of acoustic feedback before and after ANC is activated by using the l_1 controller. Control off: ———; control on: -----.

for the FULMS method and the fixed l_1 method, respectively. Although cases 5–8 contain only the results of the FULMS method and the fixed l_1 method, the l_2 method yields similar performance as the l_1 method and the corresponding results are omitted for brevity.

III. CONCLUSION

Acoustic feedback problem of the feedforward ANC structure is explored from the standpoint of control theories. Feedforward ANC systems based on l_1 and l_2 model matching algorithms have been implemented by using a digital signal processor with acoustical feedback taken into account. The ANC algorithms are derived by using the Youla's parametrization and the l_1-norm and l_2-norm vector space optimization. These methods alleviate the problems caused by the nonminimum phase (NMP) zeros of the cancellation path. Passive damping treatment was introduced to improve the robustness of the control design. It is found from the theoretical and experimental investigations that the acoustic feedback significantly complicates the control design and degrades the performance and stability of the active control system because of the modeling error involved in approximating the inverse source dynamics.

The performance of the l_1 and l_2 techniques was compared experimentally with the well-known FULMS method in the presence of acoustic feedback for three kinds of noise. Although in some cases the total attenuation appears to be comparable, these active control algorithms provide different emphasis for different frequency ranges. The experimental investigations indicate the proposed methods have potential in suppressing not only stationary noises but also transient noises, where in the latter case the FULMS algorithm is generally ineffective. This is not to mention the problem of stability and local minima with the FULMS algorithm. That makes the proposed ANC systems attractive alternative for dealing with acoustic feedback problem in industrial applications.

Despite the preliminary success, the performance of the ANC system is seriously limited by the undesirable acoustic feedback regardless of which method is used. In addition, the proposed controller is fixed in nature whose practicality will be limited in the face of system perturbations. Possibilities of different acoustic arrangements, different ANC structures with more number of transducers that are unaffected by acoustic feedback, and an adaptive version of the l_1/l_2 controller will be explored in the future study.

ACKNOWLEDGMENTS

This paper was written in memory of late Professor Anna M. Pate, Iowa State University. Special thanks also go to Professor J. S. Hu for the helpful discussions on the l_1 model matching principle. The work was supported by the National Science Council in Taiwan, Republic of China, under the Project No. NSC 83-0401-E-009-024.

[1] P. A. Nelson and S. J. Elliott, *Active Control of Sound* (Academic, New York, 1992).
[2] C. R. Fuller and A. H. von Flotow, ''Active control of sound and vibration,'' IEEE Control Systems Magazine, 9–19 (December 1995).
[3] S. M. Kuo and D. R. Morgan, *Active Noise Control Systems: Algorithms and DSP Implementations* (Wiley, New York, 1995).
[4] K. H. Eghtesadi and H. G. Leventhall, ''Active attenuation of noise. The monopole system,'' J. Acoust. Soc. Am. **71**, 608–611 (1982).
[5] D. C. Youla, H. A. Jabr, and J. J. Bongiorno, Jr., ''Modern Wiener–Hopf design of optimal controllers, Part II: The multivariable case,'' IEEE Trans. Autom. Control **AC-21**, 319–338 (1976).
[6] M. A. Dahleh and J. Diaz-Bobillo, *Control of Uncertain Systems* (Prentice-Hall, Englewood Cliffs, NJ, 1995).
[7] L. J. Eriksson, ''Development of the filtered-u algorithm for active noise control,'' J. Acoust. Soc. Am. **89**, 257–265 (1990).
[8] M. A. Dahleh and J. B. Pearson, ''l_1-optimal feedback controllers for MIMO discrete-time systems,'' IEEE Trans. Autom. Control **AC-32**, 314–322 (1987).
[9] M. A. Dahleh and J. B. Pearson, ''Optimal rejection of persistent disturbances, robust stability, and mixed sensitivity minimization,'' IEEE Trans. Autom. Control **33**, 722–731 (1988).
[10] G. Deodhare and M. Vidyasagar, ''l_1-optimality of feedback control systems: The SISO discrete-time case,'' IEEE Trans. Autom. Control **35**, 1082–1085 (1990).
[11] T. C. Tsao, ''Optimal feedforward digital tracking controller design,'' ASME J. Dynam. Syst. Measurement Control **116**, 583–592 (1994).
[12] J. C. Doyle, B. A. Francis, and A. R. Tannenbaum, *Feedback Control Theory* (Maxwell Macmillan, New York, 1992).
[13] D. G. Luenberger, *Optimization by Vector Space Methods* (Wiley, New York, 1969).
[14] W. Rudin, *Functional Analysis* (McGraw-Hill, New York, 1973).
[15] J. A. Cadzow, ''Functional analysis and the optimal control of linear discrete systems,'' Int. J. Control **17**, 481–495 (1973).
[16] K. J. Astrom, P. Hagander, and J. Sternby, ''Zeros of sampled systems,'' Automatica **20**, 31–38 (1984).
[17] R. Gueler, A. H. von Flotow, and D. W. Vos, ''Passive damping for robust feedback control of flexible structure,'' AIAA J. Guid. Control Dynam. **16**, 662–667 (1993).
[18] L. Ljung, *System Identification: Theory for the User* (Prentice-Hall, Englewood Cliffs, NJ, 1987).

Ground effect for A-weighted noise in the presence of turbulence and refraction

Keith Attenborough and Kai Ming Li
Faculty of Technology, The Open University, Milton Keynes MK7 6AA, England

(Received 6 August 1996; revised 10 March 1997; accepted 11 April 1997)

Closed-form approximations, of ground effect for continuous A-weighted noise from a point source near to finite impedance ground, obtained by Li *et al.* [J. Acoust. Soc. Am. **88**, 1170–1172 (1990)] and Makarewicz [J. Acoust. Soc. Am. **82**, 1706–1711 (1987); **88**, 1172–1175 (1990)], are extended to take account of more general models for ground impedance including revised variable-porosity, thin-layer, and rough porous ground where the roughness is small compared with incident wavelengths. The influences of both turbulence and weak atmospheric refraction within the illuminated zone are included. Approximate relationships between optimum ground effect for A-weighted noise, ground parameters, source spectrum, source height, and receiver height are predicted and explored. For low source heights and noise sources with A-weighted spectra with low peak frequencies, it is possible to indicate ground surface parameters that will give significantly higher excess attenuation than those typical of grass-covered ground. It is predicted that, in the illuminated zone, the optimum ground effect parameters are fairly robust to weak refraction effects.
 [S0001-4966(97)02608-8]

PACS numbers: 43.50.Vt, 43.28.Fp [GAD]

INTRODUCTION

Although the phenomenon of ground effect is taken into account in many prediction schemes for sound propagating from continuous broadband sources in open areas near to the ground, there have been few attempts to exploit current knowledge and understanding of the phenomenon for passive noise control. A notable exception is the design and deployment of porous road surfaces which act to reduce road/tire noise at the source and during subsequent propagation.[1]

In the absence of meteorological effects, ground effect may be predicted as a function of frequency by means of the classical theory for the propagation of sound from a point source over finite impedance ground together with either measured impedance or an impedance model.[2] Rasmussen[3] has derived an empirical approximate formula for ground effect as a function of frequency and has found it to be reliable for short-range propagation.

Most community noise problems require information about the attenuation of A-weighted noise. There are empirical schemes for predicting such attenuation for specific noise sources such as highways and railways. However, these empirical methods are not well suited for exploring ground effect optimization. In particular, empirical schemes for predicting attenuation of A-weighted noise assume that all porous ground surfaces give identical excess attenuation within the accuracy of the scheme. In fact, ground surfaces display a wide range of porosities and flow resistivities that vary over nearly three orders of magnitude. Consequently, they can give rise to a wide range of ground effects.

Although it would be possible to carry out extensive sensitivity analyses based on existing approximations for ground effect as a function of frequency and to use these as the basis for optimizing ground effect for various source spectra, the associated computation would be cumbersome and tedious and the optimization for the attenuation of A-weighted noise would have to be repeated for each given source spectrum and source–receiver geometry. On the other hand, if it were possible to devise analytical approximations for the attenuation of A-weighted sound, in which the dependencies on source spectrum, source–receiver geometry, atmospheric conditions, and ground parameters are explicit, then the optimization should be simpler and less computationally demanding.

By using exponential functions to represent A-weighting and air absorption and the Delany–Bazley single parameter model of ground impedance,[4] Makarewicz[5] derived an approximation for ground effect at relatively long range which could be used to calculate the effective flow resistivity that maximizes ground effect for a given point source–receiver geometry. Li *et al.*[6] pointed out sign errors in this work, used a two-parameter impedance model,[7] and deduced alternative closed-form results for ground effect and optimum parameters. Makarewicz[8] subsequently extended this model to take account of different source spectra and atmospheric turbulence.

In this paper, we extend this latest work to allow for a more general range of ground impedance and the effects of small scale roughness on the coherent sound field. For a given source–receiver geometry, this enables prediction of ground types that will achieve maximum excess attenuation and estimation of the magnitudes of the optimized excess attenuation. We make use of an approximation for the field in the illuminated zone under a linear sound velocity gradient to explore the extent to which the optimized ground effect at short range is affected by weak atmospheric refraction.

In Sec. I, we review the general basis for the approximations with regard to source spectra and ground effect in a homogeneous atmosphere. In Sec. II we review various

ground impedance models and introduce a new approximation for the effective impedance of rough finite impedance ground. In Sec. III, an approximate ray-based method for including weak refraction effects is outlined and justified. Subsequently we derive expressions for excess attenuation and optimum ground parameters in Sec. IV and make numerical explorations of their sensitivities to source–receiver geometry and refraction in Sec. V. Finally we offer some conclusions based on these predictions.

I. INTEGRAL EXPRESSIONS

The basis for this and previous work is the classical analytic approximation, sometimes called the Weyl–van der Pol formula, for the sound field due to a point source above an impedance plane.[2] After further approximation for near-grazing incidence, inclusion of exponential factors to allow for A-weighting, source spectrum, air absorption, and turbulence, and integration over frequency (f), an expression for the A-weighted mean-square sound pressure may be deduced.[5,8]

Hence

$$p_A^2=(\rho_0 c_0/4\pi d^2)[P_A+P_{A1}+P_{A2}], \tag{1}$$

where

$$P_A=\int_0^\infty W_0 P^{(0)}(f/f_0)^m \exp(-\mu_A f d-2\alpha(f)d)df, \tag{2a}$$

$$P_{A1}=\int_0^\infty W_0 P^{(0)}(f/f_0)^m(Q_1^2+Q_2^2)\exp(-\mu_A df - 2\alpha(f)d)df, \tag{2b}$$

$$P_{A2}=2\int_0^\infty W_0 P^{(0)}(f/f_0)^m T\{Q_1\cos(k_0\Delta r) + Q_2\sin(k_0\Delta r)\}\exp(-\mu_A df-2\alpha(f)d)df, \tag{2c}$$

$W_0=1.4676\times10^{-6}$; $f_0=1$ Hz; $\mu_A=\delta+\mu$ Hz^{-1} (or s); δ is an A-weighting parameter ($=6.1413\times10^{-4}$ s); m, μ, and $P^{(0)}$ relate to the source power spectrum such that the total A-weighted power of the source is $(P^{(0)}/\mu_A)[\Gamma(m+1)/(\mu_A f_0)^m]W$ (Γ is the gamma function); and the A-weighted power spectrum peaks at $f_m=m/\mu_A$ Hz. The air absorption parameter $\alpha(f)=a_1 f+a_2 f^2$, where a_1 and a_2 are constants ($a_2\ll a_1$) that depend on temperature and humidity, $T=\exp(-\gamma_t f^2 d)$, where $\gamma_t=\varepsilon(4\pi^{5/2}/c_0^2)\langle n^2\rangle L_0(1-\rho)$ is a turbulence parameter depending on mean-square refractive index $\langle n^2\rangle$, L_0 (m) is the largest turbulence scale, ρ is the transverse correlation, and ε is a constant ($=\frac{1}{2}$ or 1). Here Δr (m) is the path length difference between direct and ground reflected rays, d (m) is the range, and k_0 (m^{-1}) is the wave number for sound in air. In the presence of a linear sound-speed profile, and subject to certain constraints, Eqs. (2) remain valid as long as Δr in (2c) and the angle of specular reflection θ at the ground are calculated for curved ray paths (see Sec. III).

Here Q_1 and Q_2 are the real and imaginary parts of the spherical wave reflection coefficient. Assuming that A-weighting reduces any contributions from frequencies less than 250 Hz to insignificance, that $d\geqslant 10$, that $Z\cos\theta\ll 1$, where $Z=R-iX$, is the relative normal surface impedance of the ground, that θ is the angle of specular reflection, and using harmonic time dependence $\exp(i\omega t)$, where $\omega=2\pi f$, the spherical wave reflection coefficient may be approximated considerably. Hence

$$Q_1\approx -1+2R\cos\theta-\frac{2c_0}{\pi df}RX-2(R^2-X^2)\cos^2\theta, \tag{3a}$$

$$Q_2\approx -2X\cos\theta+4RX\cos^2\theta-\frac{c_0}{\pi df}(R^2-X^2), \tag{3b}$$

and

$$Q_1^2+Q_2^2\approx 1-4R\cos\theta+\frac{4c_0}{\pi df}RX+8R^2\cos^2\theta. \tag{3c}$$

Note that Eqs. (3a)–(3c) correct sign errors in the corresponding equations in Refs. 5 and 8 and misprints in Ref. 6.

It is interesting to observe that when $X>R$, the terms in (R^2-X^2) serve to increase Q_1 and Q_2, and hence the integrands in Eqs. (2b) and (2c).

An example simulation of an A-weighted source power spectrum is illustrated in Fig. 1. The crosses corresponds to $m=2.55$, $P^{(0)}=10^8$ W s, and $\mu_A=4.03\times10^{-3}$ s, giving a peak sound power level at 633 Hz, and the open boxes represent A-weighted octave band power levels deduced from the measured sound level spectrum at 152.4 m from a fixed Avon engine at a disused airfield during low wind and low turbulence conditions by correcting for spherical spreading, ground effect, air absorption, and turbulence.[9,10] In principle it should be possible to fit m, $P^{(0)}$, and μ_A, analytically, given f_m, the peak A-weighted sound power level, and P_A. However, in this instance, the values are the result of trial-and-error best-fitting.

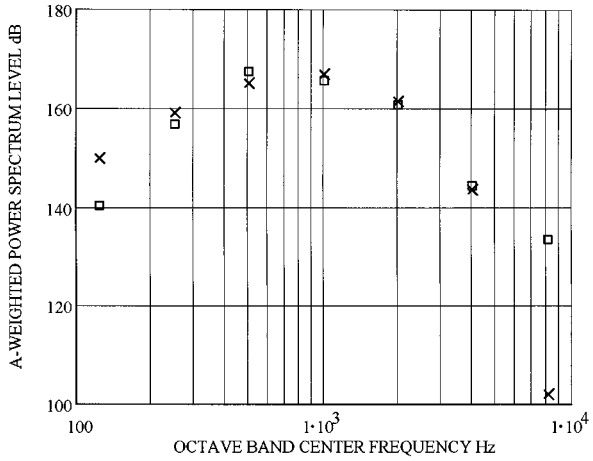


FIG. 1. Example simulation of an A-weighted octave band power spectrum (×) compared with that deduced from measurements (□) obtained at 1.2-m height with a fixed Rolls Royce Avon jet engine mounted at 2.16-m height over grassland. The simulation uses $m=2.55$, $P^{(0)}=10^8$ W s, and $\mu_A=4.03\times10^{-3}$ s.

II. APPROXIMATE MODELS FOR GROUND IMPEDANCE

Recently a rigorous derivation of the approximate surface impedance of a high-flow-resistivity rigid porous medium in which the porosity decreases exponentially with depth has been given.[11] Hence

$$Z=\frac{1-i}{\sqrt{\pi\gamma\rho_0}}\sqrt{\frac{\sigma_e}{f}}-\frac{ic_0\alpha_e}{8\pi\gamma f}, \tag{4a}$$

where $\sigma_e=s_p^2\sigma/\Omega$ and $\alpha_e=(n'+2)\alpha/\Omega$; σ, Ω, α, s_p, and n' are the flow resistivity and porosity at the surface, the rate of change of porosity, the pore shape and grain shape factors, respectively. Li *et al.*[6] used a similar but less rigorously derived model based partly on numerical simulation of the impedance of a porous material with porosity that is exponentially decreasing with depth.[7] When its flow resistivity is high, pore shape has relatively little effect on the acoustical properties of a rigid porous medium.[12] Moreover, to be consistent with the expected low-frequency limit, s_p should be set equal to 0.5 without loss of generality. An approximation for the surface impedance of a high-flow-resistivity porous medium with the porosity *increasing* exponentially with depth is given by (4a) with negative α_e. If α_e is negative, then (4a) predicts a resistance that exceeds the reactance at all frequencies. The surface impedance of a hard-backed thin high-flow-resistivity layer may be approximated by[7]

$$Z=\frac{1-i}{\sqrt{\pi\gamma\rho_0}}\sqrt{\frac{\sigma_e}{f}}-\frac{ic_0}{2\pi\gamma d_e f}, \tag{4b}$$

where $d_e=\Omega d$ and d is the layer thickness.

In general impedances of these types, both of which imply $X>R$ (as long as α_e is positive), may be written

$$Z=a(1-i)\sqrt{\frac{\sigma_e}{f}}-\frac{ib\alpha_e}{f}, \tag{4c}$$

where $a=1/\sqrt{\pi\rho_0\gamma}$ and $b=c_0/8\pi\gamma$. In the latter case, $\alpha_e=4/d_e$.

The effective normalized surface admittance of a rigid-porous boundary roughened by three-dimensional (3-D) or two-dimensional (2-D) identical bosses, the heights and spacing of which are small compared with an incident wavelength, may be approximated by[13]

$$\beta_{3D}=\beta_s+i\,\frac{2\pi f\sigma_V}{c_0}\left(\frac{3s_{3D}}{2\nu_{3D}}-1+\beta_s\frac{c_0}{c_s}\right), \tag{5a}$$

or

$$\beta_{2D}=\beta_s+i\,\frac{2\pi f\sigma_V}{c_0}\left(\frac{2s_{2D}}{\nu_{2D}}-1+\beta_s\frac{c_0}{c_s}\right), \tag{5b}$$

for 3-D and 2-D bosses, respectively, where β_s is the normalized surface impedance of the rigid-porous material surface if it were a plane, s and ν are shape and close-interaction factors, σ_V is the volume of bosses above the base-plane per unit area of surface, and c_s is the complex speed of sound within the rigid-porous material. For a high-flow-resistivity semi-infinite rigid-porous material, $\beta_s\approx\frac{1}{2}(1+i)\sqrt{\pi\rho_0\gamma f/\sigma_e}$ and $2\pi f\beta_s/c_s\approx\gamma\Omega$. Hence the rough surface effective admittance may be approximated by

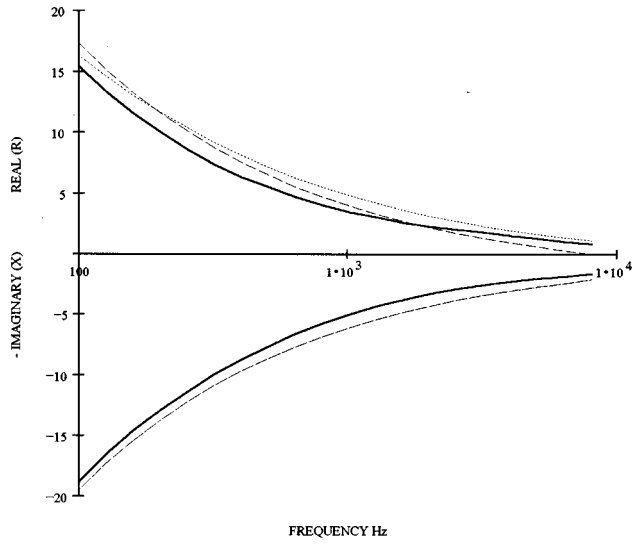


FIG. 2. Comparison of predictions of Eqs. (5a)—solid line, (5d)—dotted line, and (5e)—dashed line, for the effective impedance of a rough porous surface with $\sigma_e=200$ kPa s m^{-2} and close-packed hemispherical bosses of radius 0.01 m. Note that (5d) and (5e) give identical predictions for the reactance.

$$\beta_r\approx A(1+i)\sqrt{f}+i(C+Df), \tag{5c}$$

where $A=\frac{1}{2}\sqrt{\pi\rho_0\gamma/\sigma_e}$, $C=\sigma_v\gamma\Omega$, and $D=(2\pi\sigma_v/c_0)\times(3s_{3D}/2\nu_{3D}-1)$ or $D=(2\pi\sigma_v/c_0)(2s_{2D}/\nu_{2D}-1)$. With $C\ll A\sqrt{f}$ and $D\sqrt{f}\ll A$, the effective normalized impedance of the rough porous surface is approximated by

$$Z_r\approx\frac{1}{2A\sqrt{f}}(1-i)-\frac{1}{2A^2}(Cf+D). \tag{5d}$$

Figure 2 shows example predictions of Eqs. (5a) and (5d) for the special case of close-packed hemi-spherical bosses of radius r for which $C=\pi\gamma\Omega r/6$ and $D=(2\pi r/3c_0)[(1-\pi^2/16)/(1+\pi^2/32)]$. For close-packed identical semi-cylinders of radius r, $D=(\pi r^2/2c_0)[(1-\pi^2/12)/(1+\pi^2/12)]$. This figure demonstrates also that a valid approximation for the effective impedance of a rough porous surface of high flow resistivity is given by the relatively simple expression

$$Z_r\approx\frac{1}{2A\sqrt{f}}(1-i)-\frac{D}{A^2}, \tag{5e}$$

$c_0D/2\pi$ may be treated as an effective roughness volume per unit area.

III. EFFECTS OF WEAK REFRACTION

Apart from turbulence, meteorological effects were ignored in the Makarewicz[5,8] derivation of a closed-form expression for an A-weighted mean-square sound pressure.

However, it has been identified[14] that variations in temperature and wind with height cause the speed of sound to vary with height and lead to refraction effects. Without inclusion of refraction the applicability of the model is limited. In the past 10 years, there has been considerable progress in the prediction of sound propagation outdoors. Numerical models, such as the fast field program (FFP), the parabolic equation method (PE), etc., have been developed to compute the sound field in a complex outdoor environment. These recent developments have been summarized in Ref. 15.

In the present context, these numerical methods are not appropriate because the direct and reflected sound waves are not separated out explicitly and the evaluation of A-weighted mean-square pressure by these numerical methods demands considerable computational resources. For engineering applications and at short range, it is far more convenient to use the ray-trace approach. In the past, the ray-trace approach has been included in a heuristic modification of the Weyl–Van der Pol formula.[16,17] It has been shown recently that the resulting formulation represents the first term of an asymptotic solution of the full wave equation.[18,19] In addition, it has been demonstrated numerically that the ray-trace solution agrees reasonably well with other numerical schemes both at short ranges[16] and at long ranges[20] as long as the receiver is in the illuminated zone. In the following analysis, we shall use the ray-trace approach in deriving the expression for the A-weighted mean-square sound pressure.

It is assumed that the atmosphere is vertically stratified and that the effective sound-speed gradient can be used to replace wind and temperature gradients.[14] In order to present an analytically tractable solution, the effective sound speed, c, is assumed to vary linearly with height,

$$c(z)=c_0(1+\zeta z), \tag{6}$$

where ζ is the normalized sound velocity gradient and z is the height above ground. There is a distinct advantage for this choice because the use of a linear profile leads to a circular ray path.[14] It is relatively easy to trace the direct and reflected waves. Furthermore, the experimental data obtained by Li *et al.*[17] agree tolerably well with the theoretical predictions based a linear sound-speed profile.

To simplify our subsequent analysis, the normalized sound pressure is approximated by

$$p=\{\exp\,(ik_0\xi_1)+[Q_1+iQ_2]\,\exp\,(ik_0\xi_2)\}/4\pi d, \tag{7a}$$

where Q_1 and Q_2 are given by Eqs. (3a) and (3b), respectively. Here, in Eq. (7a), d is the horizontal separation between the source and receiver, and ξ_1 and ξ_2 are, respectively, the acoustical path lengths of the direct and reflected waves. The acoustical path lengths can be determined by[14,19]

$$\xi_1=\int_{\phi_<}^{\phi_>}\frac{d\phi}{\zeta\,\sin\,\phi}=\zeta^{-1}\,\log_e[\tan(\phi_>/2)/\tan(\phi_</2)]. \tag{7b}$$

and

$$\xi_2=\int_{\theta_<}^{\theta_>}\frac{d\theta}{\zeta\,\sin\,\theta}=\zeta^{-1}\,\log_e[\tan(\theta_>/2)\tan^2(\theta_0/2)/\tan(\theta_</2)], \tag{7c}$$

where $\phi(z)$ and $\theta(z)$ are the polar angles (measured from the positive z axis) of the direct and reflected waves. The subscripts $>$ and $<$ denote the corresponding parameters evaluated at $z_>$ and $z_<$, respectively; $z_>\equiv\max(h_s,h_r)$ and $z_<\equiv\min(h_s,h_r)$. We remark that Hidaka *et al.*[21,22] have used the travel time of the curve ray to characterize the acoustical path length. It is straightforward to show that Eqs. (7b) and (7c) can be reduced to their equations. Additionally, it is possible to show, in the limit of $\zeta\to 0$ (i.e., a homogeneous medium), the acoustical path lengths can be reduced to

$$\xi_1=\sqrt{(z_>-z_<)^2+d^2}$$

and

$$\xi_2=\sqrt{(z_>+z_<)^2+d^2},$$

which are the corresponding geometrical path lengths of the direct and reflected waves.

To allow for the computation of $\phi(z)$ and $\theta(z)$, we need to find the corresponding polar angles (ϕ_0 and θ_0) at $z=0$. Formulas for finding ϕ_0 and θ_0 in terms of sound velocity gradient are given elsewhere.[17,21] Once the polar angles are determined at $z=0$, $\phi(z)$ and $\theta(z)$ at other heights can be found by using Snell's law:

$$\sin\,\vartheta=(1+\zeta z)\sin\,\vartheta_0,$$

where $\vartheta=\phi$ or θ. Substituting these angles into Eqs. (7b) and (7c) and, in turn, into Eq. (7a), we can calculate the sound field in the presence of a linear sound velocity gradient. In particular, the path length difference, Δr, is given by

$$\Delta r=\xi_2-\xi_1. \tag{7d}$$

If we confine predictions to a horizontal separation of about 1 km and weak refraction so that the receiver is in the illuminated zone, it is reasonable to restrict the ray-trace model to a single bounce. L'Esperance *et al.*[16] point out, for downward refraction, that the additional rays will cause a discontinuity in the predicted sound level because of the inherent approximation used in the ray-trace model. It is possible to determine the critical range, r_c, where there are two additional rays in the ray-trace solution. It is tedious but straightforward to show that the critical range, for $\zeta>0$, is given by

$$r_c=\frac{\{[\sqrt{(\zeta z_>)^2+2\zeta z_>}+\sqrt{(\zeta z_<)^2+2\zeta z_<}]^{2/3}+[\sqrt{(\zeta z_>)^2+2\zeta z_>}-\sqrt{(\zeta z_<)^2+2\zeta z_<}]^{2/3}\}^{3/2}}{\zeta}. \tag{8a}$$

On the other hand, the ray-trace solution for upward refracting conditions is incorrect when the receiver is located in shadow and penumbra zones. The shadow boundary can be determined from geometrical considerations. For a given source and receiver heights, the critical range, r_c', is determined as

$$r_c' = \frac{\sqrt{(\zeta' z_>)^2 + 2\zeta' z_>} + \sqrt{\zeta'^2 z_< (2z_> - z_<) + 2\zeta' z_<}}{\zeta'}, \tag{8b}$$

where

$$\zeta' = \frac{|\zeta|}{1 - |\zeta| z_>}.$$

Using Eqs. (7a)–(7c) and following the methods detailed in Ref. 5, we can derive the A-weighted mean-square sound pressure in the presence of weak effective sound-speed gradients. The expression will be given in the next section where all other factors are included in the analysis.

IV. APPROXIMATIONS FOR EXCESS ATTENUATION

Substitution of Eqs. (4d) and (3) into (2) leads to integrals of the form

$$G_1(q) = \int_0^\infty x^q e^{-(gx + hx^2)} \cos(lx)\, dx$$

and

$$G_2(q) = \int_0^\infty x^q e^{-(gx + hx^2)} \sin(lx)\, dx,$$

where $g = \mu_A + 2a_1 d$, $h = (2a_2 + \gamma_t)d$, and $l = 0$ or $2\pi\Delta r / c_0$. These may be evaluated in closed form by means of products of gamma and parabolic cylinder functions.[5,23] If d is sufficiently large and a_2 and γ_t are sufficiently small, then $g^2 + l^2 \gg 6h$ and $l/g \ll 1$, so that[5]
for $l = 0$,

$$G_1(q) \to C(q) \approx \Gamma(q+1)/(\mu_A + 2a_1 d)^{-(q+1)}, \tag{9a}$$

for $l \neq 0$,

$$G_1(q) \approx C(q) - \tfrac{1}{2} C(q+1) l^2 - (2a_2 + \gamma_t) d C(q+2), \tag{9b}$$

and

$$G_2(q) \approx C(q+1) l - (2a_2 + \gamma_t) d C(q+3) l. \tag{9c}$$

Finally, use of these and further approximations leads to

$$p_A^2 = (P_A / 4\pi d^2)[S(\sigma_e, \alpha_e, g, m, \gamma_t, l, \theta_0, d)], \tag{10a}$$

where

$$S = B_0 d + B_1 l^2 - B_2 l \cos\theta_0 + B_3 \cos^2\theta_0 + B_4 l, \tag{10b}$$

$$B_0 = \frac{2(2a_2 + \gamma_t)}{g^2} \frac{\Gamma(m+3)}{\Gamma(m+1)}, \tag{10c}$$

$$B_1 = \frac{1}{g^2} \frac{\Gamma(m+3)}{\Gamma(m+1)}, \tag{10d}$$

$$B_2 = 4\left[\frac{\Gamma(m+\frac{3}{2})}{\Gamma(m+1)} \frac{a\sqrt{\sigma_e}}{\sqrt{g}} + b\alpha_e \right], \tag{10e}$$

$$B_3 = \frac{4g}{\Gamma(m+1)} [2\Gamma(m) a^2 \sigma_e + 2\sqrt{g}\Gamma(m - \tfrac{1}{2}) ab\sqrt{\sigma_e}\alpha_e + g\Gamma(m-1) b^2 \alpha_e^2], \tag{10f}$$

$$B_4 = \frac{2c_0 g}{\pi d \Gamma(m+1)} [2\Gamma(m - \tfrac{1}{2}) ab\alpha_e \sqrt{\sigma_e g} + g\Gamma(m-1) b^2 a_e^2]. \tag{10g}$$

The ground parameters that minimize S and hence give maximum excess attenuation, for a given geometry, are given by

$$\sqrt{\sigma_{em}} = \frac{A_2 A_4 - 2A_1 A_5}{A_4^2 - 4A_3 A_5} \tag{11a}$$

and

$$\alpha_{em} = \frac{A_1 A_4 - 2A_2 A_3}{A_4^2 - 4A_3 A_5}, \tag{11b}$$

where

$$A_1 = \frac{4a}{\sqrt{g}} \frac{\Gamma(m+\frac{3}{2})}{\Gamma(m+1)} l \cos\theta_0, \tag{11c}$$

$$A_2 = 4bl \cos\theta_0, \tag{11d}$$

$$A_3 = 8a^2 g \frac{\Gamma(m)}{\Gamma(m+1)} \cos^2\theta_0, \tag{11e}$$

$$A_4 = 8abg^{3/2} \frac{\Gamma(m-\frac{1}{2})}{\Gamma(m+1)} \left(\cos^2\theta_0 + \frac{c_0 l}{2\pi d} \right), \tag{11f}$$

and

$$A_4 = 4b^2 g^2 \frac{\Gamma(m-1)}{\Gamma(m+1)} \left(\cos^2\theta_0 + \frac{c_0 l}{2\pi d} \right). \tag{11g}$$

A similar procedure starting with the approximate rough porous surface impedance given by Eq. (5e) yields Eqs. (7a)–(7d) together with

$$B_2 = 4\left[\frac{\Gamma(m+\frac{3}{2})}{\Gamma(m+1)} \frac{a\sqrt{\sigma_e}}{\sqrt{g}} \right], \tag{12a}$$

$$B_3 = 8a^2\sigma_e \left[g \frac{\Gamma(m)}{\Gamma(m+1)} - 2aD\sqrt{\sigma_e g} \frac{\Gamma(m+\frac{1}{2})}{\Gamma(m+1)} + 2a^2\sigma_e D^2 \right], \tag{12b}$$

$$B_4 = -8a^3\sigma_e D \left[a\sigma_e D - \sqrt{\sigma_e g} \frac{\Gamma(m+\frac{1}{2})}{\Gamma(m+1)} \right]. \tag{12c}$$

The rough ground parameters that minimize S are given by

$$\sqrt{\sigma_{em}} = A_2 \Big/ 2\left(A_3 + \frac{A_4^2}{4A_5} \right) \tag{13a}$$

and

$$D_m=\frac{A_4}{2A_5\sqrt{\sigma_{em}}}, \tag{13b}$$

where

$$A_2=\frac{4a}{\sqrt{g}}\frac{\Gamma(m+\frac{3}{2})}{\Gamma(m+1)}\, l\, \cos\,\theta_0, \tag{13c}$$

$$A_3=8a^2g\,\frac{\Gamma(m)}{\Gamma(m+1)}\cos^2\,\theta_0, \tag{13d}$$

$$A_4=-16a^3\sqrt{g}\,\frac{\Gamma(m+\frac{1}{2})}{\Gamma(m+1)}\left(\cos^2\,\theta_0-\frac{c_0 l}{2\pi d}\right), \tag{13e}$$

and

$$A_4=-8a^4\left(\cos^2\,\theta_0-\frac{c_0 l}{2\pi d}\right). \tag{13f}$$

If the atmosphere is homogeneous, then $l\approx 4\pi h_r h_s/c_0 d$, $\cos\,\theta_0\approx(h_s+h_r)/d$,

$$\sqrt{\sigma_{em}}=\frac{2\pi}{ac_0g^{3/2}}\times\frac{\Gamma(m+\frac{3}{2})(h_rh_s)(h_r+h_s)}{2\Gamma(m)(h_r+h_s)^2-(\Gamma(m+\frac{1}{2}))^2\Gamma(m+1)(h_r^2+h_s^2)}, \tag{13g}$$

and

$$D_m=\frac{c_0g^2}{2\pi}\times\frac{2\Gamma(m)(h_r+h_s)^2-(\Gamma(m+\frac{1}{2}))^2\Gamma(m+1)(h_r^2+h_s^2)}{(2m+1)\Gamma(m+\frac{3}{2})(h_rh_s)(h_r+h_s)}. \tag{13h}$$

For completeness, the optimum effective flow resistivity for a semi-infinite smooth rigid-porous high-flow-resistivity surface, such that $Z=a(1-i)\sqrt{\sigma_e/f}$, under homogeneous atmosphere, is given by

$$\sigma_{em}=\frac{\pi^2}{a^2c_0^2g^3}\left(\frac{\Gamma(m+\frac{3}{2})}{\Gamma(m)}\right)^2\frac{h_r^2h_s^2}{(h_r+h_s)^2}. \tag{13i}$$

Here S may be calculated from Eq. (10b) together with Eqs. (10c)–(10g), or (10c)–(10d), and (12). Hence the excess attenuation may be calculated for optimum or other parameter values from[3,4]

$$EA=10\,\log(|S|). \tag{14}$$

V. NUMERICAL EXAMPLES AND DISCUSSION

A. Comparison with data: Homogeneous atmosphere and Avon jet engine source

With $\mu_A=4.03\times10^{-3}$ s, $P^{(0)}=10^8$ W s, and $m=2.55$, the exponentially modeled A-weighted source spectrum peaks at 633 Hz and simulates Avon jet engine noise.

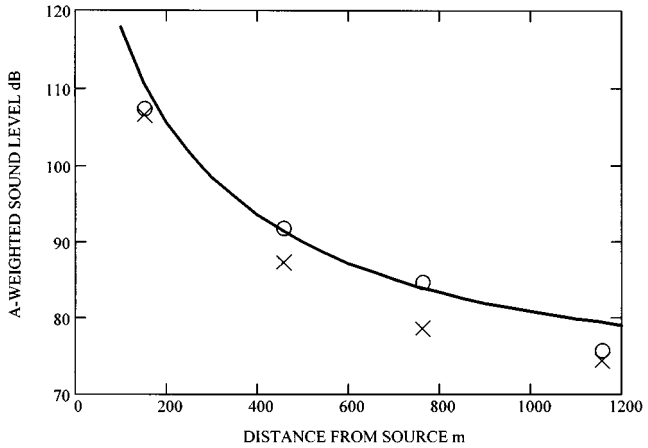


FIG. 3. Comparison of predicted (solid line) and measured sound levels (× and ○) as a function of range from a fixed Avon jet engine ($h_s=2.16$ m, $h_r=1.2$ m) under zero wind and low turbulence conditions (atmosphere characterized by $\langle n^2\rangle=10^{-8}$, $L_0=1$ m, $\rho=0$, $\varepsilon=0.5$, $a_1=5\times10^{-7}$, and $a_2=2\times10^{-12}$). The data correspond to averages over two consecutive 26-s samples.

The A-weighted sound levels predicted by the approximate theory over a variable porosity ground characterized by $\sigma_e=30$ kPa s m^{-2} and $\alpha_e=16$ m^{-1} are compared with measurements obtained under zero wind, low turbulence conditions[9] in Fig. 3. Other parameter values used in the

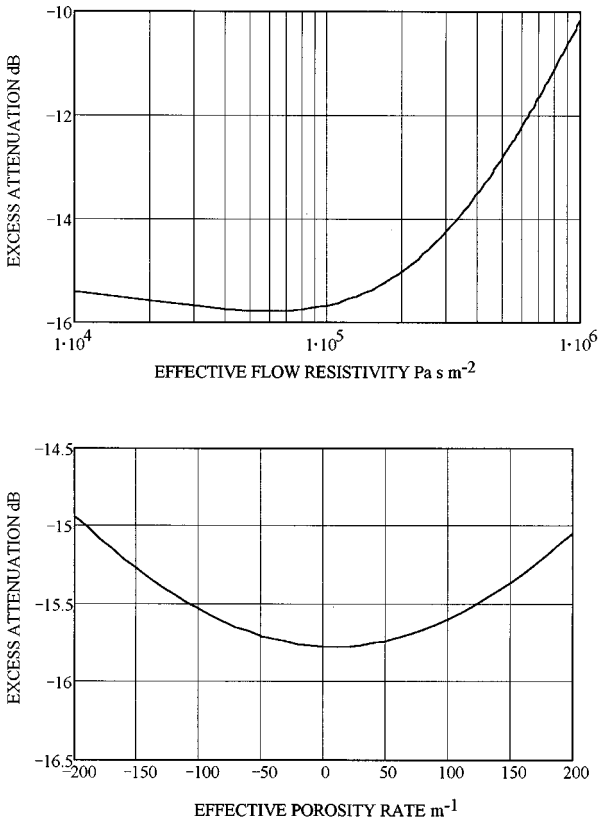


FIG. 4. Predicted variation of excess attenuation at 500 m with variable porosity ground parameters for a source at 2.16-m height, receiver at 1.2-m height, an A-weighted spectrum simulated by $m=2.55$ and $\mu_A=4.03\times10^{-3}$ s, air absorption given by $a_1=5\times10^{-7}$ and $a_2=2\times10^{-12}$, turbulence represented by $\langle n^2\rangle=10^{-8}$, $L_0=1$ m, $\rho=0$, and no refraction.

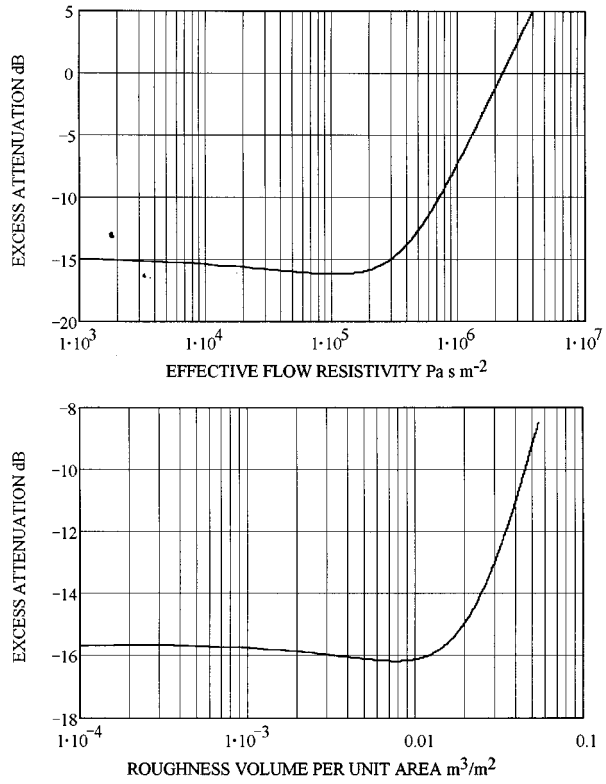


FIG. 5. Sensitivity of excess attenuation spectrum to ground parameters [rough porous ground impedance approximation, Eq. (5e)] for source with A-weighted spectrum peak frequency of 633 Hz at 2.16-m height, receiver at 1.2-m height and range 500 m, and atmospheric parameters as for Fig. 4. Variation (a) with effective flow resistivity assuming optimum roughness parameter value of 1.41×10^{-4} and (b) with roughness, assuming optimum effective flow resistivity of 107 kPa s m^{-2}.

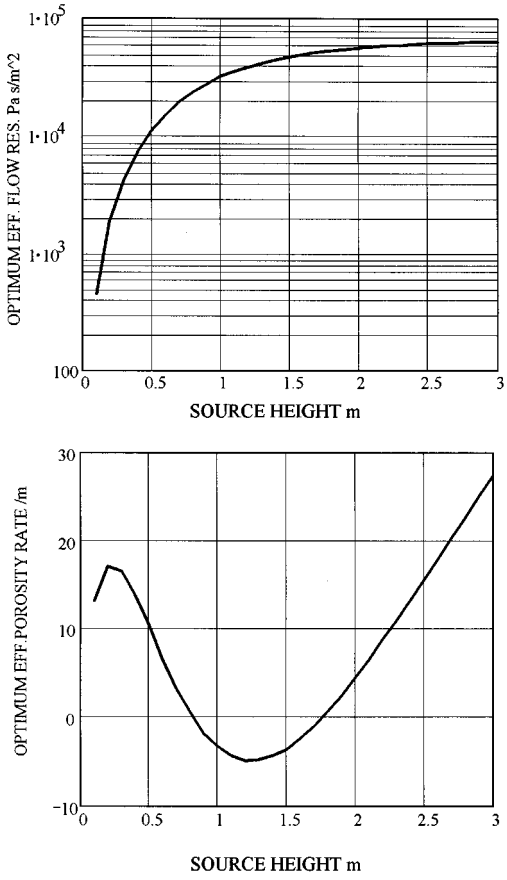


FIG. 6. Variation of predicted optimum variable porosity or thin layer parameters with source height for $h_r=1.2$ m, $d=500$ m, and other parameters as for Fig. 4.

predictions are $\langle n^2\rangle=10^{-8}$, $L_0=1$ m, $\rho=0$, $\epsilon=0.5$, $a_1=5\times10^{-7}$, and $a_2=2\times10^{-12}$.

B. Sensitivity to spectrum, source height, and distance

Adopting the variable porosity ground model, a range of 500 m, and the air absorption, turbulence, and spectrum parameter values used for Figs. 2 and 3, the optimum parameters are predicted to be $\sigma_{em}=58.8$ kPa s m^{-2} and $\alpha_{em}=7.74$ m^{-1}. However, for the given A-weighted spectrum and source height, Fig. 4 shows that the predicted excess attenuation is rather insensitive to variation in either of the ground parameters below certain values (say $\sigma_{em}<10^3$ kPa s m^{-2} and $\alpha_{em}<400$ m^{-1}). Two decades of variation in flow resistivity or large variations in α_e below these values produce only 1-dB change in predicted excess attenuation. An airport grassland[9,10] with an effective flow resistivity of approximately 30 kPa s m^{-2} and α_e of approximately 15 m^{-1} is predicted to give a ground effect fairly near to the optimum. Such a surface is predicted to give a distinctly greater excess attenuation than acoustically harder surfaces; specifically the excess attenuation is predicted to be approximately 19 dB different from that predicted for an acoustically hard surface ($+3$ dB).

Similar results are predicted, if the approximate rough porous surface impedance model is adopted. For example, the optimum rough ground parameters, for the source–receiver geometry and spectrum used for Fig. 4, are predicted to be an effective flow resistivity of 107 kPa s m^{-2} and a roughness volume per unit area of 0.0077 m. This represents soil with roughness equivalent to close-packed hemispherical bosses of radius 0.08 m. Again, as shown in Fig. 5, the excess attenuation is predicted to be insensitive to effective flow resistivity below a value of 100 kPa s m^{-2}. However, increasing roughness beyond the optimum is predicted to have a significantly detrimental effect on A-weighted excess attenuation.

The optimum parameters are predicted to be very sensitive to source height and spectrum. This is illustrated in Figs. 6 and 7 for an Avon engine spectrum.

The corresponding optimum excess attenuations are predicted to be very sensitive to turbulence. This is demonstrated in Figs. 8 and 9. The predictions for both ground types are similar.

Clearly the most useful optimizations of the excess attenuation of A-weighted noise due to ground effect are predicted to occur for low source heights and low peak frequencies in the A-weighted source spectrum. It should be noted that the predictions for low flow resistivities will be somewhat suspect since high flow resistivity impedance models have been assumed.

Considerable increase in predicted excess attenuation over that for typical grassland is predicted to be possible for

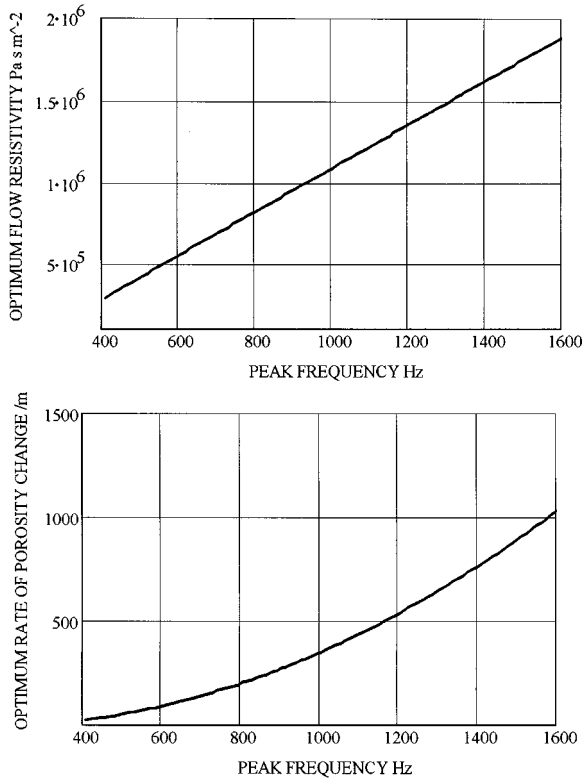


FIG. 7. Variation of predicted optimum variable porosity or thin layer parameters with A-weighted spectrum peak frequency: $h_s = 2.16$ m, $h_r = 1.2$ m, and other parameters as for Fig. 4.

low source heights if the flow resistivity is low and either the porosity is near constant with depth or the roughness is fairly large. This suggests potential noise reduction from plowing or disking ground since either reduces flow resistivity while increasing roughness. The effect of ground treatment on flow resistivity may be temporary as a consequence of weathering.

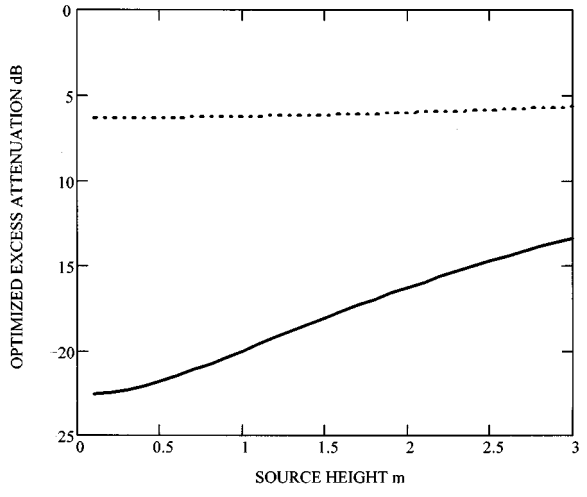


FIG. 8. Variation of predicted optimum excess attenuation with source height for variable porosity or thin layer ground with source height: $h_r = 1.2$ m, $d = 500$ m, and other parameters as for Fig. 4. Dotted and solid lines represents predictions with strong and weak turbulence ($\langle n^2 \rangle = 10^{-6}$ and 10^{-8}, respectively).

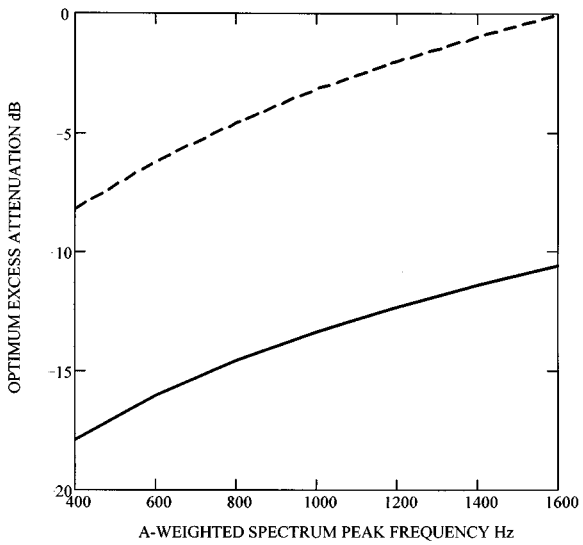


FIG. 9. Predicted variation of optimum excess attenuation with A-weighted spectrum peak frequency for $h_s = 2.16$ m, $h_r = 1.2$ m, $d = 500$ m, other parameters as for Fig. 4, and for two values of turbulence: $\langle n^2 \rangle = 10^{-8}$ (solid) and ($\langle n^2 \rangle = 10^{-6}$ (broken).

C. Variation with distance

The optimum ground parameters are predicted to depend on distance from the source. Figure 10, for weak turbulence

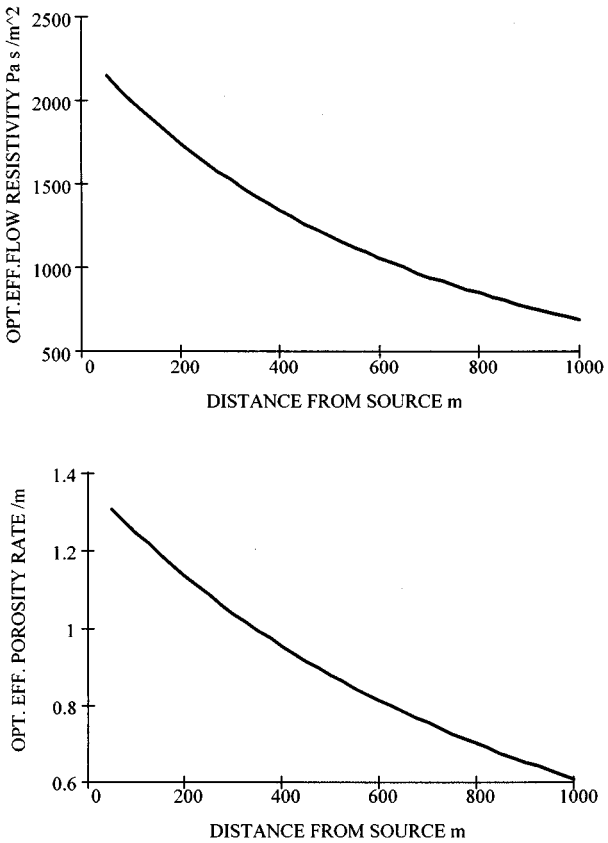


FIG. 10. Predicted optimum ground parameters as a function of distance for an A-weighted spectrum characterized by $m = 1.5$, $\mu_A = 2 \times 10^{-3}$ s, atmospheric parameters as for Fig. 4, source height 0.1 m, and receiver height 1.2 m.

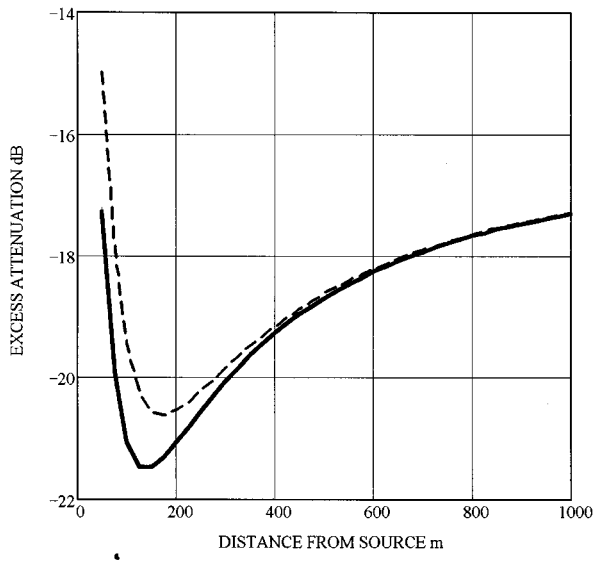


FIG. 11. Predicted optimized ground effect (solid line) for source characterized by $m=2$, $\mu_A=2\times10^{-3}$ Hz^{-1}, source height 0.1 m and receiver height 1.2 m, compared with ground effect predicted over optimized ground type ($\sigma_e=3$ kPa s m^{-2}, $\alpha_e=3$ m^{-1}, dotted but indistinguishable from solid line on graph) and over an airfield ground ($\sigma_e=30$ kPa s m^{-2}, $\alpha_e=16$ m^{-1}, dashed line). Air absorption is represented by $a_1=0.5\times10^{-6}$ and $a_2=2\times10^{-12}$, turbulence by $\langle n^2\rangle=10^{-8}$ and $L_0=1$ m.

and a 0.1-m-high source characterized by an A-weighted spectrum peaking at 750 Hz, illustrates the prediction that the optimum parameters change towards acoustically softer values as distance from source increases. However, the dependence is relatively slight and the predicted excess attenuation over ground with constant parameters average within the optimum range is similar to the optimized excess attenuation. This is illustrated in Fig. 11.

D. Effects of refraction

Figures 12–15, for weak turbulence, low source height, and low-frequency spectrum peak (300 Hz), show that the predicted optimum ground parameters represent increasingly hard ground under increasing downward refraction conditions and increasingly soft ground under increasing upward

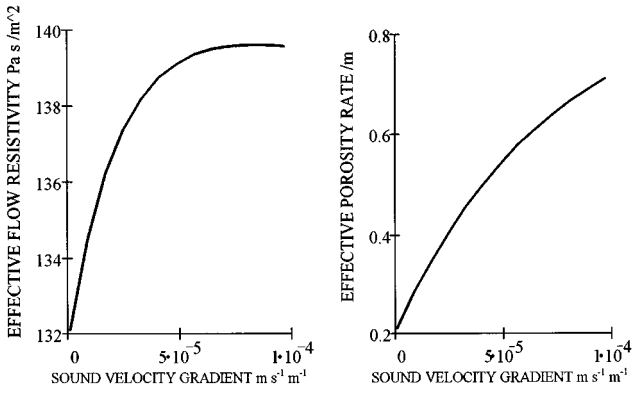


FIG. 12. Variation of optimum variable porosity ground parameters with downward refracting sound velocity gradient for $m=1.5$, $\mu_A=5\times10^{-3}$ Hz^{-1}; Source height 0.1 m, receiver height 1.2 m, range 200 m, atmospheric parameters as for Fig. 11.

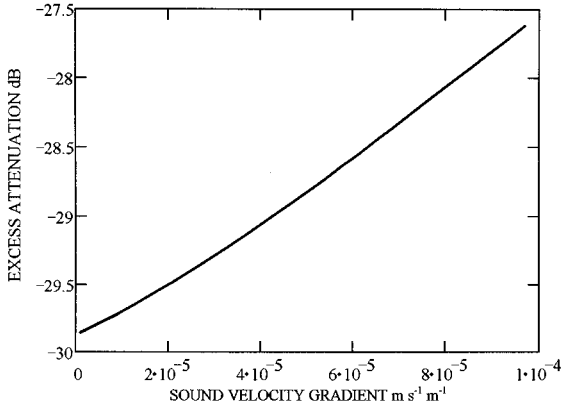


FIG. 13. Variation of optimum excess attenuation with downward refracting sound velocity gradient for source spectrum, absorption, turbulence, and geometry specified for Fig. 12.

refraction conditions. The predicted optimum excess attenuations decrease as downward refraction increases and increase as upward refraction increases. Nevertheless, the dependence is relatively slight for the given conditions, so that the predicted optimum ground effect is fairly robust to weak refraction.

VI. CONCLUDING REMARKS

By means of approximations of the classical theory for a point source above an impedance surface and approximate effective impedance models it is possible to derive closed-form relationships between ground effect related excess attenuation, parameters relating to the A-weighted source power spectrum and the ground and source and receiver

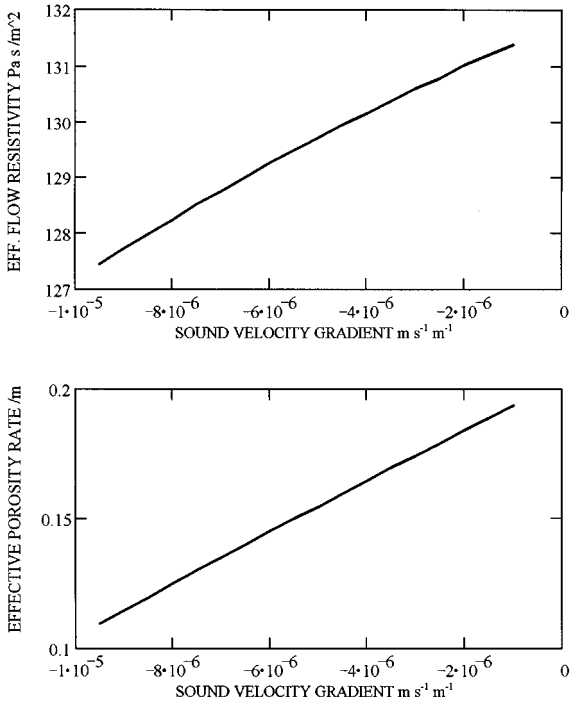


FIG. 14. Predicted effect of upward refraction on optimum variable porosity ground parameters. Geometry, spectrum, absorption, and turbulence parameters have the values used for Figs. 12 and 13.

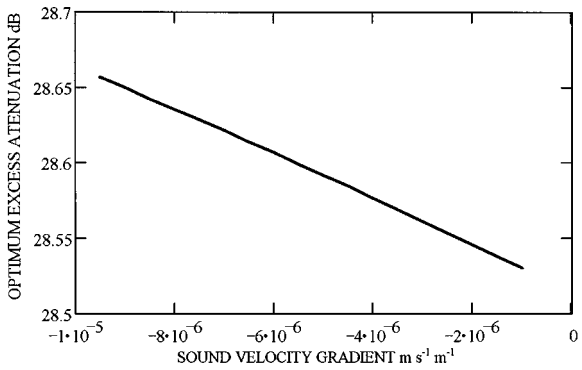


FIG. 15. Predicted effect of upward refraction on optimum excess attenuation. Geometry, spectrum, absorption, and turbulence parameters have the values used for Figs. 12 and 13.

heights. Although the resulting predictions are approximate they suggest practical possibilities for optimizing excess attenuation by means of controlling the ground characteristics. The greatest opportunities for such optimization are predicted for sources with low-frequency A-weighted spectral peaks and at ranges of less than 200 m under weak turbulence conditions and require rather acoustically soft ground. This conclusion is consistent with the known effects of appreciable snow cover. The influence of atmospheric refraction has been included by means of a ray-based approximation valid at short range in the illuminated zone. The suggested optimizations are fairly robust to weak atmospheric refraction.

The optimum attenuations are predicted to be much higher for sources close to the ground. Since it may not be practicable to lower source heights for many noise sources, an alternative strategy for exploiting ground effect is to raise the ground, perhaps through landscaping. This possibility will be the subject of future work.

ACKNOWLEDGMENTS

This work was supported in part by EPSRC (UK) Grant ref. J42052 and BBSRC (UK) Grant ref. CTE 02649. The authors and the paper have benefitted also from detailed and helpful reviews.

[1] K. Attenborough and C. Howorth, in *Proceedings of the International Tire-Road Noise Conference* (Swedish National Board for Technical Development 1990), Vol. 1, pp. 171–199.

[2] K. Attenborough, ''Review of ground effects on outdoor propagation from continuous broadband sources,'' Appl. Acoust. **25**, 289–319 (1988).

[3] K. B. Rasmussen, ''Approximate formulae for short distance outdoor propagation,'' Appl. Acoust. **29**, 313–324 (1990).

[4] M. E. Delany and E. N. Bazley, ''Acoustical properties of fibrous absorbent materials,'' Appl. Acoust. **3**, 105–116 (1970).

[5] R. Makarewicz, ''Near-grazing propagation above a soft ground,'' J. Acoust. Soc. Am. **82**, 1706–1711 (1987).

[6] K. M. Li, K. Attenborough, and N. W. Heap, ''Comment on 'Near-grazing propagation above soft ground,' '' J. Acoust. Soc. Am. **88**, 1170–1172 (1990).

[7] K. Attenborough, ''Ground parameter information for propagation modeling,'' J. Acoust. Soc. Am. **92**, 418–427 (1992).

[8] R. Makarewicz, J. Acoust. Soc. Am. **88**, 1172–1175 (1990) (reply to Ref. 6 above).

[9] K. Attenborough, K. M. Li, and S. Taherzadeh, ''Propagation from a broad-band source over grassland: Comparison of data and models,'' Proc. INTERNOISE 95 **1**, 319–322 (1995).

[10] ESDU Data Item No. 940355, ''The correction of measured noise spectra for the effects of ground reflection'' (ESDU International, London, 1994).

[11] R. Raspet and J. M. Sabatier, ''The surface impedance of grounds with exponential porosity profiles,'' J. Acoust. Soc. Am. **99**, 147–152 (1996).

[12] K. Attenborough, ''Models for the acoustical properties of air-saturated granular media,'' Acta Acust. **1**, 213–226 (1993).

[13] K. Attenborough and S. Taherzadeh, ''Propagation from a point source over a rough finite impedance boundary,'' J. Acoust. Soc. Am. **98**, 1717–1722 (1995).

[14] I. Rudnick, ''Propagation of sound in open air,'' in *Handbook of Noise Control*, edited by C. M. Harris (MacGraw-Hill, New York, 1957), Chap. 3, pp. 3:1–3:17.

[15] K. Attenborough, S. Taherzadeh, H. E. Bass, X. Di, R. Raspet, G. R. Becker, A. Güdesen, A. Chrestman, G. A. Daigle, A. L'Espérance, Y. Gabillet, K. Gilbert, Y. L. Li, M. J. White, P. Naz, J. M. Noble, and H. A. J. M. van Hoof, ''Benchmark cases for outdoor sound propagation models,'' J. Acoust. Soc. Am. **97**, 173–191 (1995).

[16] A. L'Esperance, P. Herzog, G. A. Daigle, and J. R. Nicolas, ''Heuristic model for outdoor sound propagation based on an extension of geometrical ray theory in the case of a linear sound speed profile,'' Appl. Acoust. **37**, 111–139 (1992).

[17] K. M. Li, K. Attenborough, and N. W. Heap, ''Source height determination by ground effect inversion in the presence of a sound velocity gradient,'' J. Sound Vib. **145**, 111–128 (1991).

[18] K. M. Li, ''On the validity of the heuristic ray-trace based modification to the Weyl–Van der Pol formula,'' J. Acoust. Soc. Am. **93**, 1727–1735 (1993).

[19] K. M. Li, ''Propagation of sound above an impedance plane in a downward refracting atmosphere,'' J. Acoust. Soc. Am. **99**, 746–754 (1996).

[20] R. Raspet, A. L'Esperance, and G. A. Daigle, ''The effect of realistic ground impedance on the accuracy of ray tracing,'' J. Acoust. Soc. Am. **97**, 154–158 (1995).

[21] T. Hidaka, K. Kageyama, and S. Masuda, J. Acoust. Soc. Jpn. (E) **6**, 117–125 (1985).

[22] In Ref. 21 the angles are measured from the horizontal axis whereas, in this paper, the angle is measured from the vertical (z) axis.

[23] I. Gradshteyn and I. Ryzhik, *Tables of Integrals, Series and Products* (Academic, New York, 1980), p. 496, 3.953.1 and 2.

Three-dimensional noise field directionality estimation from single-line towed array data

Ronald A. Wagstaff and J. Newcomb
Naval Research Laboratory, Code 7176, Stennis Space Center, Mississippi 39529-5004

(Received 5 July 1995; accepted for publication 22 April 1997)

For many reasons, the three-dimensional (3-D) arrival structure of the undersea ambient noise field is of interest to the research and development community. One reason is that the arrival structure can be used to estimate the beam noise of an array, which may be required to estimate the performance of the array as an operational Navy asset or a scientific measurement tool. Another reason is that there are clues inherent in the vertical arrival structure that relate to the nature of the acoustic propagation along the azimuths of the noise sources. Similarly, there are also clues in the horizontal arrival structure of the undersea ambient noise field that relate to the azimuthal distribution of the noise sources. Both of these classes of clues are important in the verification and validation of undersea ambient noise models. The ideal measurement tool to measure the 3-D arrival structure of the noise field is a high-resolution volumetric array sonar system. Unfortunately, such a system is not generally available. However, towed horizontal line arrays are available and can be used, even though they are far from ideal. The beam patterns of a line array are conically symmetric about the axis of the array. When the axis of a horizontal line array is tilted slightly from an elevation angle of 0°, the vertical character of the beam cones can be used to discriminate between various vertical arrival angles. Combined with measurements made on several different towed array headings, an estimate of the 3-D directionality of the noise field can be obtained. An algorithm that uses such a technique on single-line towed array data to generate an estimate of the 3-D directionality of the noise field is described herein. Some results using measured data from a recent noise measurement exercise are presented and discussed. [S0001-4966(97)03508-X]

PACS numbers: 43.60.Gk, 43.60.Pt, 43.30.Nb, 43.30.Wi [JLK]

INTRODUCTION

In order to design sonar systems which operate with maximum effectiveness, it is necessary to understand as much as possible about the undersea ambient noise. Two characteristics of the noise that have been of particular concern in the past are its vertical and horizontal directionalities. Urick[1] gives an excellent summary of undersea ambient noise directionality measurements, and Etter[2] describes computer models that predict the directional characteristics of the noise for various environmental and noise source input parameters. Unfortunately, those noise models are limited by our ability to accurately calculate the acoustic environment, to specify the radiation characteristics of the noise sources, and to predict their spatial distributions. Our capabilities in some of these areas have improved considerably in the past few years, but our capabilities in other areas have not advanced as much. For example, the development of the parabolic equation (PE) solution to the wave equation has greatly increased our ability to accurately calculate propagation in complex acoustic environments.[3] This, in turn, improves our ability to model the ambient noise in those environments. On the other hand, the complete radiation characteristics of ships are still not well understood. For example, the radiation directivity patterns of the noise caused by the cavitating screws are known to vary with elevation and azimuth angle and with screw depth, but a quantitative description for general classes of ships is not presently available. Hence, there are still enough uncertainties in various aspects of the prediction process, that the development of noise models will continue to be an ongoing process. As the noise models improve, the requirements on measured results to verify and validate the models will be more specific, and the nature of the processing of the noise measurement data to satisfy those specific needs will be much more demanding.

In the past, conventional two-dimensional (2-D) results from vertical arrays, which integrate over azimuth, and from towed horizontal line arrays, which integrate over vertical arrival angle, have been adequate for noise model validation. This was because the greatest need has been predicting the performance of single-line array sonar systems. However, multiple-line array sonar systems and even large multidimension array sonar systems are being considered and tested.[4,5] Predicting the performance of these complex sonar systems and other equally complex systems of the future requires as much knowledge about the 3-D arrival structure of the ambient noise as possible. This knowledge must be acquired by using data measured with the sonar systems that are in common use at the present. The bulk of those measurements will, of necessity, be made with line arrays, including single-line towed arrays.

The horizontal single-line towed array is considered an excellent tool for measuring the 2-D horizontal directionality of the ambient noise field. A common method of measurement is to tow the array on several different courses and use the beam noise measurements thus acquired to resolve the left–right ambiguities in the beam noise response patterns.

The data thus acquired usually are processed by an algorithm that ignores the vertical dimension and confines itself entirely to the horizontal plane or azimuth dimension. Such an algorithm yields an estimate of the vertically integrated (or collapsed on the horizontal plane) horizontal directionality of the ambient noise. Several techniques have been devised for resolving the left–right ambiguities and producing 2-D estimates of the noise field horizontal directionality. A few examples are presented and discussed in Refs. 6–11. All of the techniques known to the authors, except the one in Ref. 11, ignore the vertical arrival structure of the noise. Furthermore, the technique in Ref. 11 only uses the vertical arrival structure to obtain a better estimate of the 2-D horizontal structure of the noise field. None of the previous techniques known to the authors treat the horizontal and vertical dimensions as independent regimes.

The noise field 3-D estimation technique discussed herein can be considered similar in concept to the two 2-D techniques discussed in Refs. 10 and 11. It differs in implementation from the two 2-D techniques in two critical ways. First, it keeps the elevation (vertical) and azimuth (horizontal) dimensions independent. Second, it utilizes the entire 3-D beam noise response patterns of the line array. This means that, with as little as a one-degree vertical tilt on the towed array, the measured vertical ambiguities can be resolved. A tilted towed array condition is easily met, since it is seldom that the towed array is exactly horizontal during measurements. Thus by maintaining the vertical and horizontal independence, the technique discussed herein can produce an estimate of the complete 3-D directionality of the ambient noise field.

The next section provides a brief background discussion. The section that follows describes the 3-D noise field directionality estimation technique. That section is followed by a discussion of some results, and then observations and conclusions are given.

I. BACKGROUND

Knowledge of the 3-D structure of the undersea ambient noise is of value for a number of reasons. From a scientific perspective, the vertical arrival structure of a signal measured by a vertical array tells something about the average acoustic propagation conditions along the propagation path (or azimuth) of the signal. Analogously, the azimuthal arrival structure measured by a horizontal array tells something about the azimuthal distribution of the noise sources. Even when the results of these two types of arrays are used together (incoherently) there will be blind spots in the noise sphere (or 3-D space) that cannot be estimated with a 2-D analysis. Such blind spots could, in fact, be measured by a 2-D or 3-D array. Hence, a technique that can estimate the 3-D structure of the noise field from more readily available and more easily acquired single-line towed array data would provide information now to enable the sonar design engineer to design future multidimensional (2-D or 3-D) arrays as well as to predict their performance.

The vertical line array is well studied for measuring the vertical directionality of the ambient noise. However, its conical beam responses are formed about a vertical axis, and they integrate the arrivals from all azimuth angles that are received at a given elevation angle. Unfortunately, the beams cannot distinguish one arrival at a given azimuth from another at a different azimuth that has the same elevation angle. The horizontal line array has a different limitation. Its conical beams are formed along a horizontal axis and have both azimuth and elevation angle dependence. Some techniques have been devised to utilize the beam noise measured, while the horizontal array is on different headings, to estimate the horizontal directionality of the ambient noise.[6–11] However, each of those techniques, except the technique in Ref. 11, assumes that the ambient noise is concentrated in the horizontal plane (i.e., an elevation angle of 0° which is equivalent to no vertical arrival structure). The technique in Ref. 11 differs substantially from the others, in that it utilizes an estimate of the shape of the azimuthally averaged ambient noise vertical directionality, similar to that provided by a vertical line array, to improve the estimate of the horizontal directionality. However, it does not produce a 3-D estimate of the ambient noise, just a better estimate of the 2-D horizontal directionality. In the process, it reduces the error caused by ignoring the vertical arrival structure of the ambient noise. Hence, the technique in Ref. 11 is a substantial improvement over the other techniques, but it falls short of the desired 3-D estimate of the noise field.

Acquiring a measurement of the 3-D structure of the noise is not a simple task. Generally, a multidimensional array would be required, but those arrays are usually expensive to construct and difficult to deploy. On the other hand, the single-line towed array is extremely easy to deploy, sometimes requiring less than 1 h for deployment and usually not much more than that to achieve a stable towing condition. However, it has the limitation of conical beams rather than the ''ideal searchlight'' beams of a volumetric array. The technique presented in this paper overcomes, to some degree, the conical limitation of the single-line towed array beams and provides an estimate of the 3-D structure of the ambient noise field. Such an estimate can be used in the design and in the performance prediction of simple and complex sonar systems, including systems that have multidimensional arrays.

II. APPROACH

It is assumed that the ambient noise $N(\theta,\phi,t)$ can be expressed as

$$N(\theta,\phi,t)=n(\theta,\phi)+\zeta(\theta,\phi,t)+\epsilon(\theta,\phi,t), \tag{1}$$

where θ is the spherical angle in azimuth, ϕ is the spherical angle in elevation, t is time, $n(\theta,\phi)$ is the pseudo-stationary background noise field directionality, $\zeta(\theta,\phi,t)$ is the time-dependent component of the noise due to fluctuations in acoustic propagation, noise source movement, changes in noise source levels, etc., and $\epsilon(\theta,\phi,t)$ is the error introduced in the measurements by the towship noise, array nonlinearities, flow noise, system faults, etc.

The pseudo-stationary background directionality term $n(\theta,\phi)$ is the part of Eq. (1) that the technique presented herein is attempting to estimate. It is the term which is considered herein to be more characteristic of what most of the

present noise models attempt to predict. With sufficient temporal averaging, the influence of the fluctuating component $\zeta(\theta,\phi,t)$ can be considered to be negligible. However, if the performance of a signal processor is to be predicted (not the case here), the importance of $\zeta(\theta,\phi,t)$ cannot be over emphasized. Furthermore, it is assumed that the error component $\epsilon(\theta,\phi,t)$ can be made acceptably small by array grooming and appropriate error discrimination processing techniques.

The data from which $n(\theta,\phi)$ is estimated (see Ref. 10) is the set of beam-output noise intensities measured by the ith beam while on the jth array heading. This can be represented as

$$r_{i,j}=\frac{1}{T}\int_0^T \frac{1}{2\pi}\,dt\int_0^{2\pi}\frac{1}{2}d\theta\int_{-\pi/2}^{\pi/2}N(\theta,\phi,t)\times b_i(\theta-\gamma_j,\phi)\cos\,\phi\,d\phi, \tag{2}$$

where T is the measurement time interval, $b_i(\theta,\phi)$ is the ith beam response pattern of I beams, and γ_j is the jth array heading of J array headings.

With temporal averaging, array grooming, and appropriate error discrimination processing techniques, Eq. (2) reduces to

$$r_{i,j}\approx\frac{1}{4\pi}\int_0^{2\pi}d\theta\int_{-\pi/2}^{\pi}n(\theta,\phi)b_i(\theta-\gamma_j,\phi)\cos\,\phi\,d\phi. \tag{3}$$

Here, $n(\theta,\phi)$ is obtained indirectly by first forming an initial guess $\hat{n}_0(\theta,\phi)$ (which could simply be a constant value for all θ and ϕ), substituting $\hat{n}_0(\theta,\phi)$ in place of $n(\theta,\phi)$ in Eq. (3), and performing the integration to get an initial estimate $\hat{r}_{i,j,0}$ for each beam and array heading. $\hat{n}_0(\theta,\phi)$ and $\hat{r}_{i,j,0}$ are the initial values for the more general $\hat{n}_k(\theta,\phi)$ and $\hat{r}_{i,j,k}$, respectively, where the iteration number k is set to zero (i.e., $k=0$).

In general, the differences for a given iteration

$$\Delta_{i,j,k}=r_{i,j}-r_{i,j,k} \tag{4}$$

are accumulated for all θ and ϕ within the response footprint of the ith beam on the jth array heading for all I beams and J array headings. The $\Delta_{i,j,k}$ are divided by 2, individually mapped to, and distributed over each θ and ϕ. This result is then added to $\hat{n}_k(\theta,\phi)$ for each θ and ϕ to create a new estimate $\hat{n}_{k+1}(\theta,\phi)$ which is again used in Eq. (3). This iterative process is continued until the $\Delta_{i,j,k}$ are less than some arbitrary small value.

The iterative technique just described is not limited to towed array data, although much emphasis is placed on its use. In fact, data from any array can be used in the 3-D noise field estimation algorithm, so long as the corresponding 3-D beam response pattern of the array is available. For example, an improved estimate of the 3-D arrival structure of the noise field could be achieved by supplementing towed array data with data acquired concurrently from a vertical line array. This will be demonstrated by the results presented herein. The vertical array may have different beam patterns and a different number of beams than the towed array, but such differences are inherently covered by the generic nature of Eq. (3). As far as the 3-D estimation algorithm is concerned, the vertical array data is simply another data set with different beam patterns. In the extreme, even the data from a 3-D array could be used. In such a case, the 3-D algorithm would still deconvolve the beam patterns from the data and produce an improved deconvolved estimate of the 3-D noise field.

Figure 1 illustrates the 3-D response of four different beams. The beam noise response spheres as well as the 3-D ambient noise field are represented as planes (similar to a ''flat earth'' plot or a Mercator projection) with the elevation angles ϕ plotted along the vertical axis, and the azimuth angles θ plotted along the horizontal axis. Figure 1(a) shows the response of a beam near aft end fire for a horizontal line array towed on a heading of 50°. The multicolored region gives the beam response level of the main lobe. In an attempt to avoid excessive clutter in the plot, the complete beam response pattern including the sidelobes has not been presented. However, the complete patterns are used in the algorithm and in generating the results discussed below. Figure 1(b) shows a similar response for the broadside beam of the same array. The beam response in Fig. 1(b) cover all azimuth angles at $\pm 90°$ because it contains both the ''north and south'' poles. When the line array is tilted relative to the horizontal plane, as in Fig. 1(c) for a vertical tilt angle of 30°, the response pattern no longer contains the north and south poles. Instead, it forms a serpentine pattern about the equator (0° elevation angle), and covers a region in the vertical that is approximately equal to 180° minus twice the tilt angle. Figure 1(d) shows the beam response for a beam steered 56° from forward end fire, when the array has a vertical tilt angle of 30°. As a general rule, the beam patterns in the rectangular display of the θ-ϕ domain are rather complicated. However, they can be calculated once and stored for use over and over again in the iterative process during the generation of the estimate $\hat{n}_k(\theta,\phi)$.

The set of beam response patterns is calculated in spherical space by transforming the intersection of the conical pattern with a unit sphere by the following transformation.[11]

$$\beta=\cos^{-1}\{\cos\,\theta\,\cos\,\phi\,\cos\,\alpha+\sin[\cos^{-1}(\cos\,\theta\,\cos\,\phi)]\times\sin\,\alpha\,\sin[\tan^{-1}(\sin\,\phi/\cos\,\phi\,\sin\,\theta)]\}, \tag{5}$$

where β is the conical angle of a differential element on the surface of a unit sphere, and α is the tilt angle of the array. Figure 2 illustrates the geometry of Eq. (5) (Ref. 11 gives a more complete discussion).

The generation of the 3-D noise field directionality estimate can be summarized as follows: assume a 3-D noise field [e.g., isotropic, $\hat{n}_0(\theta,\phi)=C$]; transform the spherical noise field to the conical field of the array [e.g., use Eq. (5)]; obtain beam noise estimates $\hat{r}_{i,j,0}$ using $\hat{n}_0(\theta,\phi)$; compare resulting beam noise estimates $\hat{r}_{i,j,0}$ with measured beam noise $r_{i,j}$; modify $\hat{n}_0(\theta,\phi)$ at all cells in the θ-ϕ domain, within the coverage footprint of the beam, according to the differences $\Delta_{i,j,0}=r_{i,j}-\hat{r}_{i,j,0}$ to get a new estimate $\hat{n}_1(\theta,\phi)$; and repeat the above procedure until acceptable agreement is achieved (the $\Delta_{i,j,k}$ are less than some arbitrary small value).

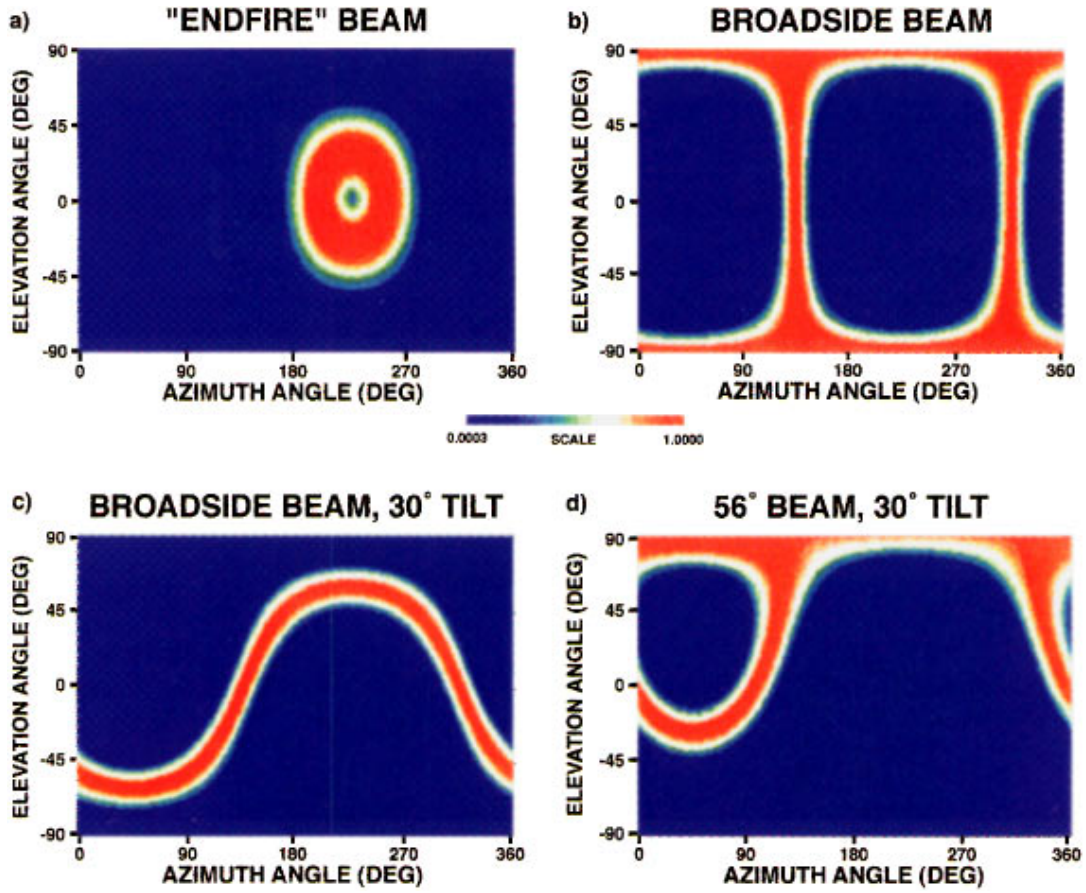


FIG. 1. Three-dimensional beam response patterns for a horizontal line array with no tilt [plates (a) and (b)] and with 30° tilt [plates (c) and (d)].

III. RESULTS

Ambient noise measurements were made by the SACLANT Undersea Research Centre's (SACLANTCEN) horizontal single-line towed array at three locations in a shallow water area using the SACLANTCEN ship (the R/V ALLIANCE) as the towing platform during August and September 1993. In addition, ambient noise measurements were made concurrently at these locations with the SACLANTCEN vertical line array. The results discussed herein were obtained using data from two of the shallow water sites designated site 7 and site 1.

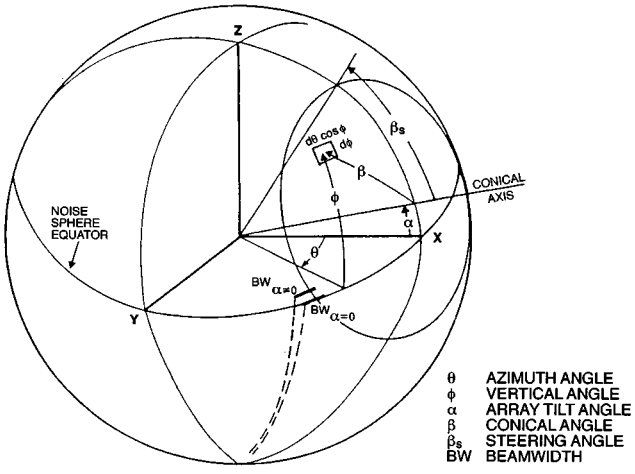


FIG. 2. Coordinate system representation.

Single-line towed array beam noise data were collected at each site on at least nine different array headings (i.e., nine legs). The headings of the legs were at appropriate intervals to insure a reasonable sampling in azimuth. The towed array generally had a tilt from the horizontal (0° elevation angle) of approximately 1° to 2°. The tilt of the towed array is used by the algorithm to help resolve the up–down ambiguities of the conical beams. In addition, a vertical line array was deployed and data were measured at each of the sites concurrently with at least five of the horizontal line array legs. Data from the vertical array were also used independently to help determine the vertical arrival structure of the noise field for comparison with the result from the 3-D calculation. The measured hydrophone data from both arrays were spectrum analyzed and 64 beams per array heading were formed by a frequency domain beamformer[12] which creates beams nonlinearly spaced in real azimuth space. Results for each array and each leg were then assessed for acceptable data quality.[13] Those data that passed the quality assessment were then processed with the 3-D algorithm as discussed above.

Figure 3 shows the results for 100 Hz at site 7 that were obtained from the 3-D algorithm by using only the towed horizontal line array data. The upper left-hand plot in Fig. 3 is the 3-D representation of the spatially averaged and temporally smoothed noise field. The vertical axis is a represen-

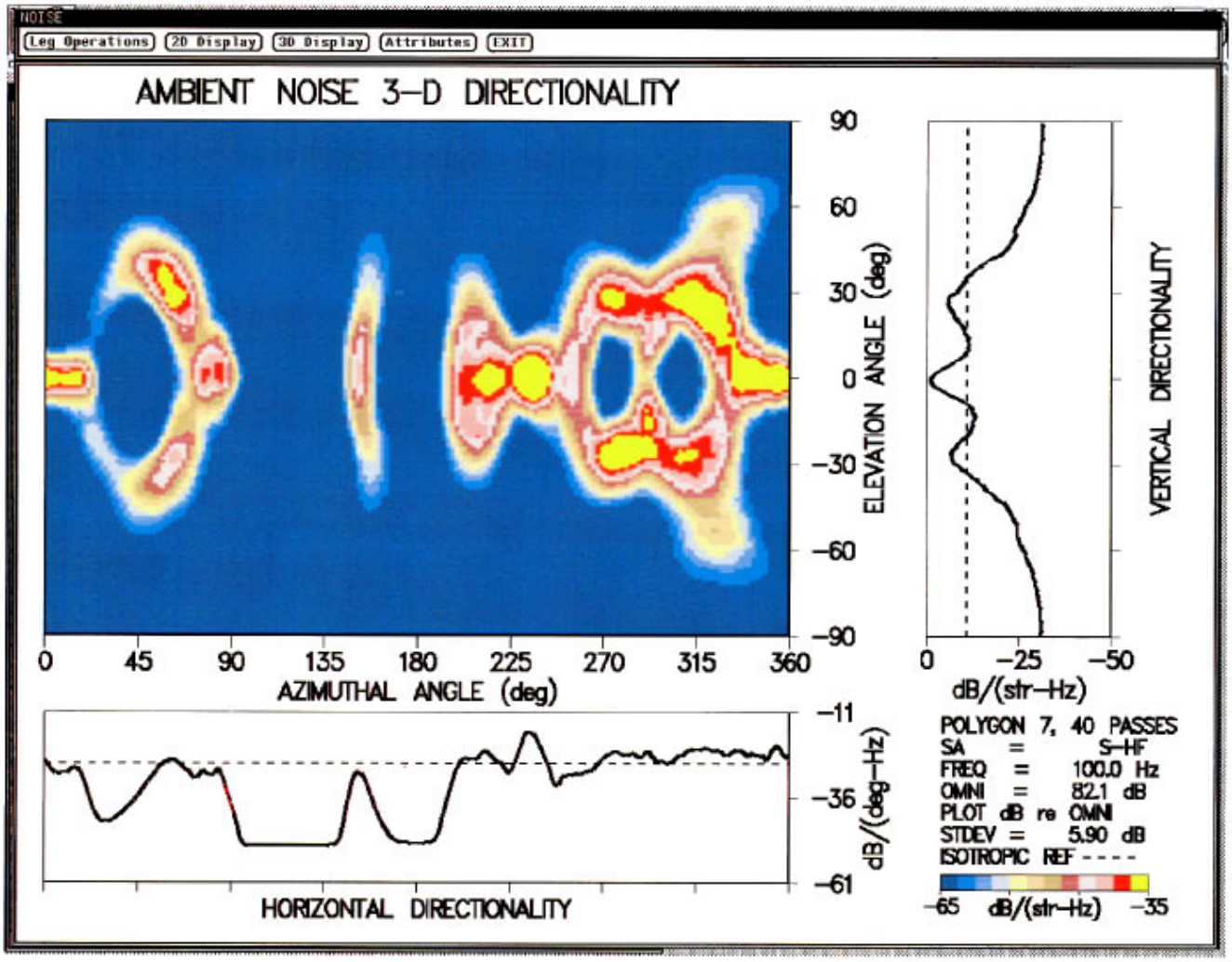


FIG. 3. Ambient noise 3-D directionality (top left) at site 7 with the corresponding 2-D horizontal directionality (bottom) and vertical directionality (top right). Only data from a towed line array were used.

tation of the angular range of the vertical (or elevation) angle ϕ. Vertically up is 90° and vertically down is −90°. The horizontal axis is a representation of azimuth angle from 0° to 360° of the compass (north to north). The magnitudes of the noises that are plotted relative to the omnidirectional level in decibels are given by the color bar at the lower right-hand corner of Fig. 3.

The plot in the upper right-hand corner of Fig. 3 is the estimate of the vertical directionality of the noise field represented by the noise surface plot in the upper left-hand corner. The estimates of the vertical directionality in the upper right-hand corner are plotted relative to the omnidirectional level in decibels. The solid black curve represents the average levels obtained by summing the noise power along all azimuths at a given elevation in the surface plot. The curve is an estimate of what a vertical line array would have measured if it had been at approximately the same measurement location and depth as the towed array and with approximately the same temporal averaging. The black vertical dashed line is an isotropic average power reference line. Deviation of the vertical directionality curve from the isotropic reference line illustrates the degree of anisotropy of the vertical directionality curve.

The plot at the bottom of Fig. 3 gives the spatially smoothed and temporally averaged 2-D azimuthal directionality of the ambient noise field (solid black curve) plotted relative to the omnidirectional level in dB. This curve is obtained in a manner similar to the black curve in the vertical directionality plot by summing the power over all vertical angels for each azimuth angle. It is this curve that most closely resembles the 2-D estimates of the ambient noise field that other techniques such as those in Refs. 6–11 attempt to generate. The black horizontal dashed line in this plot is the cylindrically isotropic average power level reference curve. Deviation from that reference curve is a measure of the anisotropy of the noise field in azimuth.

Summary information and some pertinent parameters have bene included at the bottom right-hand corner of Fig. 3. The levels in all of the plots are relative to the omnidirectional noise level (OMNI) in dB. The frequency (FREQ) is 100.0 Hz and the standard deviation of the estimate (STDEV) is 5.90 dB. The standard deviation is based on the differences between the measured beam noise levels for each leg and the corresponding beam noise levels obtained from the estimate. A value of 5.90 dB is a relatively low standard deviation for the estimation process and indicates that the

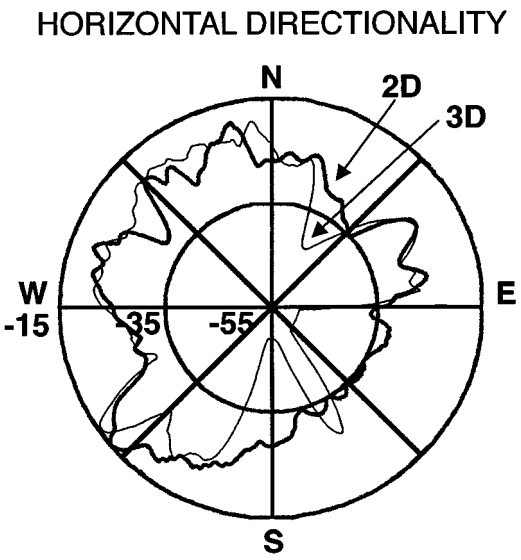


FIG. 4. Ambient noise 2-D directionalities for 100 Hz at site 7 obtained from the 3-D algorithm (thin curve) and from the 2-D algorithm from Ref. 10. Noise levels are relative to the per-horizontal-degree omnidirectional reference curve (dashed circle).

estimate of the 3-D directionality can be accepted with a high degree of confidence.

One way of evaluating the 3-D algorithm is to compare it with the 2-D algorithm[10] which has enjoyed wide acceptance for many years. The comparison can be done on the basis of the directionality that is obtained when the noise is summed over all elevation angles of the surface plots (solid black curve at the lower left-hand corner of Fig. 3). This process eliminates the vertical dimension and produces a 2-D noise directionality curve that can be directly compared to the corresponding curve produced by the 2-D algorithm. The comparison is made simple if this curve is presented in a polar plot.

Figure 4 permits the comparison of the horizontal directionalities for 100 Hz at site 7 that were obtained from the previous 2-D algorithm[10] (thick curve) and from the 3-D algorithm. The two omnidirectional levels that were obtained from each directionality estimate differed no more than a fraction of a decibel. However, there are significant differences between the two estimates along some azimuths, and there are remarkably good agreements between the two estimates along large regions of azimuth. There are four regions of azimuth where the 3-D estimate is much lower than the 2-D estimate (i.e., near 30°, 120°, 180°, and 225°). On the other hand, the 2-D algorithm gave a lower estimate along three small regions of azimuth (i.e., near 305°, 330°, and 355°). It is not possible from these results, or from similar results for the other sites, to quantify the magnitude of the agreement and disagreement in a statistically significant manner. However, an ''eye ball'' evaluation suggests that the 2-D and the 3-D algorithms are generally producing similar 2-D results. The 3-D algorithm can be expected to achieve lower levels in some areas of azimuth because it is able to partition the noise more realistically in three dimensions than the 2-D algorithm can partition the noise in two dimensions. After all, noise is a three-dimensional quantity, and the 3-D algorithm is more closely tied to a realistic 3-D arrival structure and 3-D response pattern of the array. Hence when there are disagreements between the two algorithms, the 3-D estimate would have the higher level of confidence associated with it. The limited number of simulations that have been performed tend to support this contention.

The ambient noise 3-D directionality surface plot in Fig. 3 is interesting for several reasons. First, both the left–right and up–down ambiguities have been resolved. Second, although the ambient noise, for the most part, is concentrated within about 30° of the horizontal, there are azimuths at which the majority of the noise does not arrive at or near 0° elevation angle, but arrives in concentrations around $\pm 30°$ elevation angles. There are also azimuths at which very little noise arrives, regardless of elevation angle. Third, the algorithm (and plot) provides an estimate of the 3-D noise field that was obtained from data acquired by a towed horizontal line array, an asset that is common to the inventories of many research establishments or can be rented from the petroleum exploration community. Furthermore, the plot provides the results from which the performance of much more complex and expensive multidimensional arrays can be estimated. Finally, the results provide insight into the 3-D arrival structure of the noise field that must otherwise come from measurements by multidimensional arrays. At the present time, measurements with multidimensional arrays are not realistic expectations. Given the current projections for undersea research, this may not be a reasonable expectation for many years to come.

Figure 5 show the results for 100 Hz at site 7 obtained from the 3-D algorithm using data from both the towed horizontal array and the vertical line array. Comparison of the results in Figs. 3 and 5 show that the inclusion of the vertical array data at this site had a tendency to concentrate the noise more toward 0° elevation angle. In all other respects, the two figures are very similar.

The two colored dashed curves in the horizontal directionality plot in the bottom of Fig. 5 are the azimuthal directionalities (appropriately normalized) of slices at two different elevation angles in the surface plot. The red curve is a slice at 0° elevation angle and the blue curve is a slice at $+30°$ elevation angle.

The two colored dashed curves in the vertical directionality plot in the right-hand plot of Fig. 5 are vertical slices across the noise surface plot showing the vertical directionality at tow different azimuths. One slice (the light blue vertical curve) is where the noise has one of the many maxima near the horizontal, and the other (the pink vertical curve) is where the noise has a minimum near the horizontal and maxima near $\pm 15°$ elevation angle. The difference between the extremes in the vertical directionality curves is more than 30 dB. This difference could not be measured by either a horizontal or a vertical array separately or combined without the 3-D algorithm. Furthermore, before development of the 3-D algorithm, a directionality estimate similar to the surface plot in Fig. 5 would have been possible only with data from a 2-D or 3-D array.

Specific features of the noise shown in Fig. 5 can be correlated to the noise source distribution and the environ-

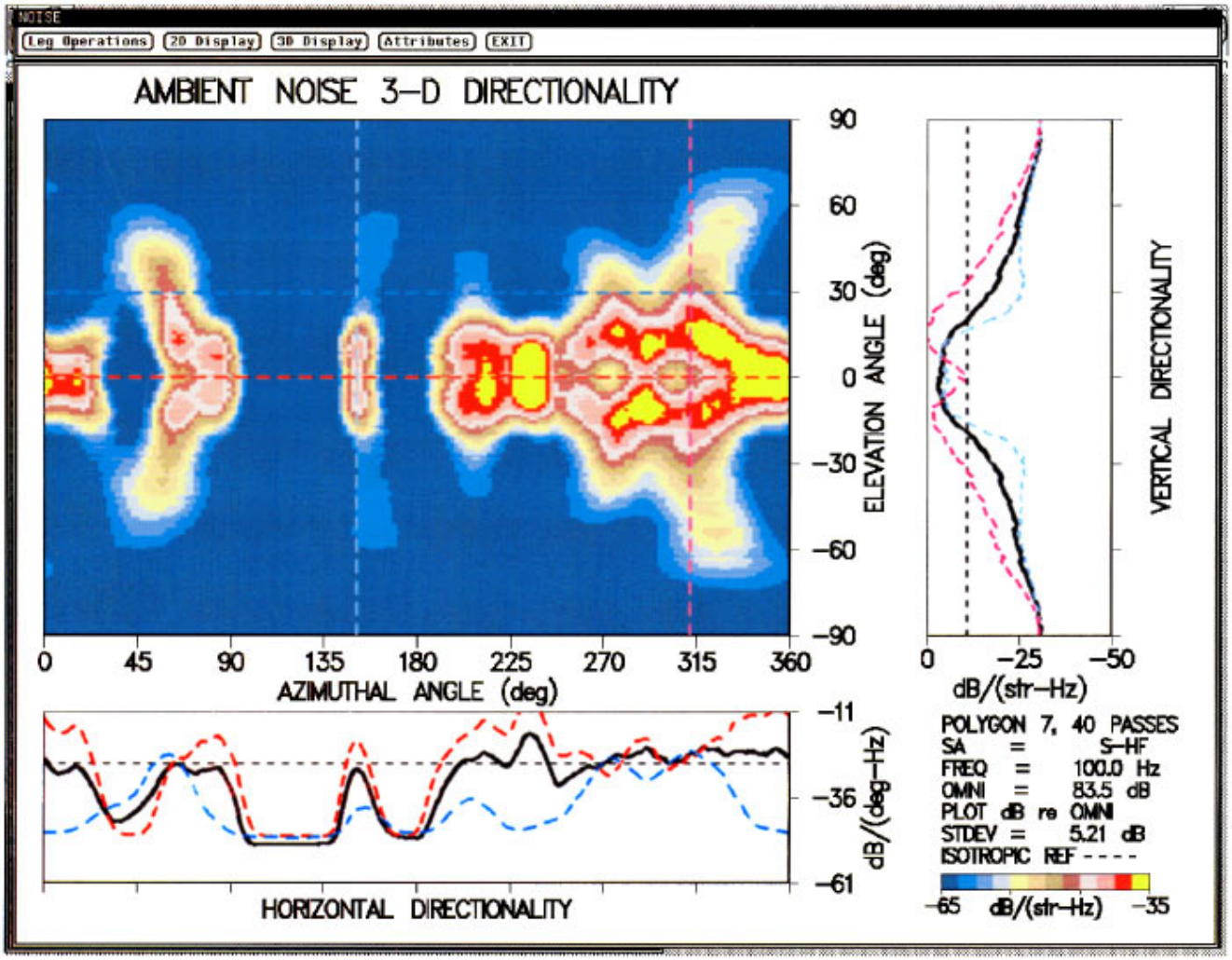


FIG. 5. Ambient noise 3-D directionality plot for 100 Hz at site 7 using data from both towed line and vertical line arrays were used. Dashed curves correspond to vertical and horizontal cuts across the surface plot.

ment. The noise arriving near 0° elevation angle along an azimuth of approximately 150° comes from the major northern access for coastal steamers to a nearby port. Noise generated closer to the port and at the southern access to the plot is blocked by small islands as well as the relatively shallow bathymetry in these areas (160°–185°). The noise arriving near 0° elevation angle from azimuths between approximately 185° and 245° is most likely distant shipping noise that is propagating to the measurement area from the open basin. Along azimuths of 245° to 330°, the noise is arriving from a chain of nearby islands. Propagation conditions along these azimuths are not conducive to long-range propagation or down slope conversion of the noise to vertical arrival angles near 0°.[14] The bathymetry in this area, when compared to the bathymetry along the southern side of the measurement area, has a gentler slope and is not as highly reflective. Furthermore, these paths are along the slope, which gives rise to a propagation path which curves away from site 7. Thus only nearby shipping can contribute from these azimuths. This may in part explain the reason for the noise concentrations at ±20° elevation angles rather than at near 0° elevation angle. The noise arriving near 0° elevation angle along azimuths between approximately 330° and 25° comes from the direction of the routes for coastal steamers in that area. The azimuths between 25° and 60° includes up slope propagation paths.

Figure 6 illustrates the results for 100 Hz at site 1 which were obtained from the 3-D algorithm using data from both the towed horizontal array and the vertical line array. Site 1 is outside the island chain directly to the north of site 7 on a relatively shallow continental shelf. It is evident from Fig. 6 that the noise at site 1 is concentrated between vertical arrival angles of ±30°. Similar to site 7 in Fig. 5, there are azimuths which have the noise concentrated near 0° elevation angle, azimuths which have noise concentrations at high vertical angels, and azimuths which have very little noise at any vertical arrival angle. The broad region of noise around north (345°–25°) is believed to be due to distant ships. The quiet region between 25° and 45° can be attributed to adverse propagation along a steep slope. As was discussed earlier, this type of propagation condition can curve the noise away from a straight-line path to the measurement site and reduce the noise considerably along affected azimuths. The noise arriving along azimuths between 45° and 110° is from a region that includes shipping accesses to various ports and routes for coastal steamers. The region of noise from azi-

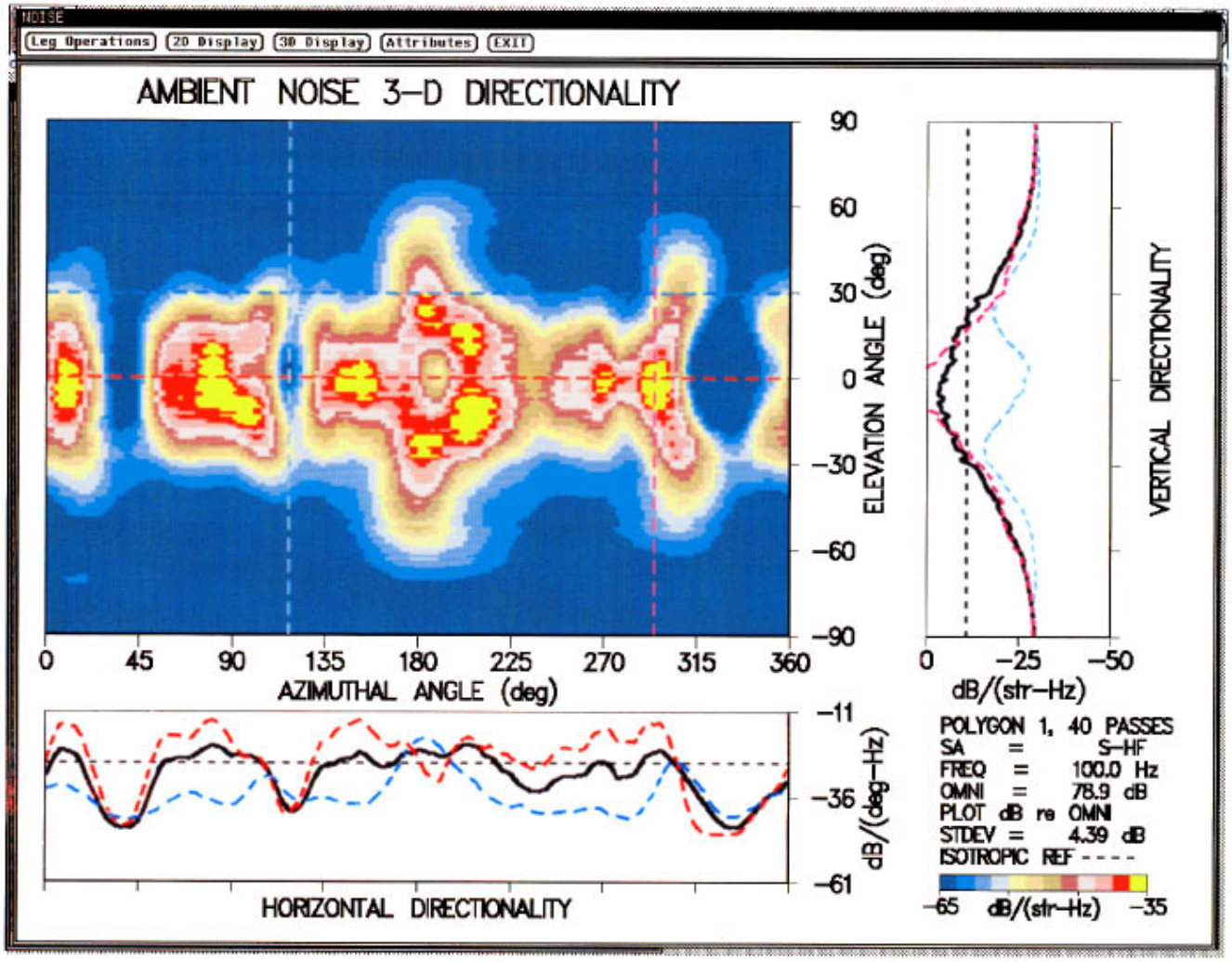


FIG. 6. Ambient noise 3-D directionality plot for 100 Hz at site 1 using data from both towed line and vertical line arrays were used. Dashed curves correspond to vertical and horizontal cuts across the surface plot.

muths between 110° and 160° is in the direction of the seaward side of an island chain. This area has a very steep downward sloping bottom, which can convert the vertical arrival angles of the acoustic energy into elevation angles that are very near to 0° elevation angle.[14] The noise along azimuths from 160° to 225° is concentrated at vertical arrival angles that are close to ±20°. This feature is indicative of propagation conditions that do not encourage conversion of energy from higher vertical arrival angles into lower elevation angles. Indeed, the bathymetry along these azimuths rarely goes below 200 m for at least 100 nmi. In addition, this region is dominated near site 1 by a claystone low loss bottom. The azimuths between 225° and 315° receive noise near 0° elevation angle, which is consistent with distant shipping. Finally, it is believed that the quite area along azimuths between 315° and 345° occurs because no distant shipping noise is expected from these directions.

IV. OBSERVATIONS AND CONCLUSIONS

The pseudo-stationary ambient noise field 3-D arrival structure can be significantly different and more complex than either its vertical or horizontal directionality. Both are significantly degraded by the spatial smoothing that results from a 2-D horizontal or vertical analysis. However, the 3-D nature of the method presented herein is able to resolve the arrival structure of the noise simultaneously in the vertical angel and the horizontal angle. In doing so, the regions of low-level noise are well identified and separated in elevation and azimuth angles from the regions of high-level noise. The resulting 3-D noise field arrival structure estimate can be used to predict the noise response of any array, whether it is a relatively simple vertical or horizontal line array, a more complicated tilted line array, a planar array, or any multidimensional array.

The pseudo-stationary ambient noise field 3-D arrival structure can also be correlated to the propagation environment to gain insight into the source distributions and propagation mechanisms active during the measurement. This was performed for two sites in a shallow water environment to identify important propagation parameters and features. Site 7 was not exposed to distant shipping from most azimuths, but did receive noise near 0° elevation angle on many azimuths. This was most likely due to environmental conditions which supported down-slope conversion of the acoustic energy into vertical angles closer to 0° elevation angle. The

noises at site 1 were more concentrated near 0° elevation angle and could be associated with distant shipping. At both sites, however, the same down-slope conversion conditions existed to allow for noise arriving near 0° elevation angle from directions in which distant shipping was not possible. For both sites, there were situations in which the distant shipping noise contributions did not appear in the 3-D estimate of the noise field even though distant ships were known to be present. These were from azimuths that were along, rather than across, a steep slope that would cause the sound propagation paths to bend away from a straight-line path to the measurement site and cause an absence of noise near 0° elevation angle.

ACKNOWLEDGMENTS

Partial funding support of the 6.2 Full Spectrum Noise Project at the Naval Research Laboratory is gratefully acknowledged. The authors are grateful for the opportunity to participate in the SACLANT Undersea Reserach Centre's (SACLANTCEN) measurement program and for the opportunity to use SACLANTCEN measurement assets, including the research vessel R/V ALLIANCE, the measurement instrumentation including the acoustic arrays, the data acquisition systems, and even measured data. The cooperation and helpfulness of all personnel aboard the R/V ALLIANCE was excellent. We are especially indebted to Mr. Tuncay Akal, the Chief Scientist aboard the R/V ALLIANCE and Dr. Reginald Hollett for playing key roles in facilitating the acquisition of the data. Appreciation is also expressed to Stephanie F. Kooney, a fellow NRL researcher, Karen Dudley, a co-op student, and David Lynch, of Alliant Techsystems, Inc. (San Diego, CA) for their valuable assistance in acquiring, transcribing, and processing the data. We also express our appreciation to them for keeping the data processing system functioning after unexpected changes were required to overcome some unpredictable problems that arose during the data transcription and processing.

[1] R. J. Urick, ''Ambient noise in the sea,'' Undersea Warfare Technology Office, Naval Sea Systems Command, 1984.
[2] P. C. Etter, *Underwater Acoustic Modeling* (Elsevier, New York, 1991), Chap. 7.
[3] F. D. Tappert, ''The Parabolic Approximation Method,'' in *Wave Propagation and Underwater Acoustics*, edited by J. B. Keller and J. S. Papadakis, Lec. Notes Phys., Vol. 70 (Springer-Verlag, New York, 1977).
[4] P. D. Koenig, ''Transient response of multidimensional arrays,'' Naval Underwater Systems Center Technical Report No. TR 7193, 1984.
[5] V. C. Anderson, ''Nonstationary and nonuniform oceanic background in a high-gain acoustic array,'' J. Acoust. Soc. Am. **67**, 1170–1179 (1982).
[6] A. H. Nuttall, ''Estimation of noise directionality spectrum,'' Naval Underwater Systems Center Technical Report No. TR 4345, 1972.
[7] A. H. Nuttall, ''Resolving the directional ambiguities of a line array of hydrophones,'' Naval Underwater Systems Center Technical Report No. TR 4385, 1972.
[8] J. H. Wilson, ''Ambient-noise horizontal directionality measurements with linear arrays,'' J. Acoust. Soc. Am. **60**, 955–960 (1976).
[9] E. M. Wilson, ''Directional noise measurements with line arrays,'' Admiralty Research Laboratory, Teddington, Middlesex, Report No. ARL/M/N16, 1973.
[10] R. A. Wagstaff, ''Iterative technique for ambient-noise horizontal-directionality estimation from towed line-array data,'' J. Acoust. Soc. Am. **63**, 863–869 (1978).
[11] R. A. Wagstaff, ''Horizontal directionality estimation considering array tilt and noise field vertical arrival structure,'' J. Acoust. Soc. Am. **67**, 1287–1294 (1980).
[12] J. R. Williams, ''Fast beam-forming algorithm,'' J. Acoust. Soc. Am. **44**, 1454–1455 (1968).
[13] R. A. Wagstaff, ''A computerized system for assessing towed array sonar functionality and detecting faults,'' IEEE J. Ocean Eng. **18**, 529–542 (1993).
[14] R. A. Wagstaff, ''Low-frequency ambient noise in the deep sound channel—the missing component,'' J. Acoust. Soc. Am. **69**, 1009–1014 (1981).

Effects of middle ear pressure on otoacoustic emission measures

Ming Zhang and Paul J. Abbas
Department of Speech Pathology and Audiology and Department of Otolaryngology—Head and Neck Surgery, University of Iowa, Iowa City, Iowa 52242

(Received 18 June 1996; revised 7 April 1997; accepted 8 April 1997)

An animal surgical model was established to manipulate pressure in the middle ear. The $2f_1-f_2$ cochlear microphonic distortion product (CMDP) and distortion product otoacoustic emission (DPOAE) were simultaneously measured in order to analyze the effects of middle ear pressure on sound transmission. The frequency ratio of the two stimulus tones f_2/f_1 was held at 1.2 and the level ratio L_1/L_2 was 10 dB. The effect of middle ear pressure on forward transmission was determined from the CMDP data. The effect on backward transmission was determined from the DPOAE data after the effect on forward transmission was subtracted out. The results showed a similar effect on forward and backward transmission. In general, negative pressure had a greater effect than positive pressure. Positive pressures of 10 and 20-cm H_2O affected transmission for low-frequency stimuli ($f_2=1620$ and 2680 Hz) but had little effect for high-frequency stimuli ($f_2=6980$ and 10 250 Hz). Negative pressures of -2.5 to -10-cm H_2O affected transmission across all frequencies tested. The effect at low frequencies is hypothesized to be related to tympanic membrane stiffness. The effect of negative pressure at high frequencies may be related to changes in the incudostapedial joint. © *1997 Acoustical Society of America.* [S0001-4966(97)02008-0]

PACS numbers: 43.64.Bt, 43.64.Ha, 43.64.Jb [RDF]

INTRODUCTION

Otoacoustic emissions (OAEs) can be utilized in the diagnosis of hearing disorders (Lonsbury-Martin *et al.*, 1993), particularly for subjects who cannot be tested behaviorally, such as very young children or those who are not conscious (Norton, 1993). Many factors may affect the measurement of the OAEs and consequently their interpretation (Kemp *et al.*, 1990). One factor is pressure change in the middle ear, which can be produced by a gas exchange mechanism (Trine *et al.*, 1993). Such a pressure change is common in middle ear disease. The incidence of middle ear disease is higher in young subjects, the population for which otoacoustic emission (OAE) measures may be particularly applicable (Fria *et al.*, 1985; Howie, 1975; Teele *et al.*, 1980). The present study attempts to develop a better understanding of the effects of abnormal middle ear pressure on signal transmission by measuring OAEs and cochlear microphonic potentials where middle ear pressure is changed under experimental control. The effect on transmission may be dependent on both frequency and level of the stimulus. In this study, we have evaluated the effect of middle ear pressure on both forward and backward transmission across a range of stimulus frequency, using a guinea pig model.

I. METHODS

Twenty pigmented female adult guinea pigs (Hartle and Janes Co.) were used as experimental subjects. The procedures were approved by the Animal Care Committee at the University of Iowa. The animal was considered acceptable if the external ear canal and tympanic membrane appeared normal. Surgery and data collection took place within a sound-treated booth.

General anesthesia was induced by an intramuscular injection of ketamine (14.7 mg/kg), xylazine (3.67 mg/kg), and acepromazine (0.37 mg/kg). Anesthesia level was monitored by paw-pinch reflex and maintained as needed throughout the experiment. A single bolus of atropine sulfate (0.05 mg/kg) was administered to reduce mucosal secretion and the risk of tracheal obstruction (Harvey and Walberg, 1987). Local anesthesia with 1% Lidocaine was injected posterior to the pinna. Throughout the experiment, the body temperature was maintained by a warm pad placed under the animal and insulating drapes.

An incision was made along the posterior sulcus of the pinna to expose the lateral wall of the mastoid bulla. Two holes were drilled through the lateral wall of the mastoid bulla. One hole was used for placement of the cochlear microphonic electrode at the round window. The other hole was used for insertion of a plastic tube (0.5-mm i.d.) connected to a pressure source, which included a syringe to supply air pressure, a U-tube manometer to monitor pressure, and a rubber-skin drum to compensate for the changes in air volume. In this experiment, pressures were set at -10, -5, -2.5, 0, 10, and 20 cm H_2O.

The middle ear was inspected under a surgical microscope through these two holes to assure that there was no pathology. A solder ball electrode measuring 1 mm in diameter and insulated with a silicon tube was placed on the posterior edge of the round window through one hole. Cochlear microphonic recordings were made differentially using the round window ball electrode (+), a subdermal electrode rostral to the ear canal (−), and a subdermal electrode caudal to the ear canal (GND).

A microphone-earphone assembly tip was sealed into the external ear canal about 4 mm from the tympanic membrane. The acoustic stimulation, sound-pressure measurement, and cochlear microphonic recordings were generated by an Ariel-DSP board (Ariel Corporation) and controlled by

EMAV v3.1 software (Neely & Liu, Boys Town National Research Hospital), both of which were installed in a Gateway-2000 486-DX computer. The computer program controlled signal generation, signal averaging, and spectral analysis. Two tones (f_1 and f_2) were output on separate digital-to-analog (D/A) channels, through transformers, to two ER-2 insert earphones connected to the assembly tip. Either the amplified voltage output from a ER-10 microphone or the amplified cochlear microphonic potential was fed into an analog-to-digital (A/D) converter. The program measured and plotted the recorded amplitude of each primary and the amplitude of the distortion products. Five frequency pairs of primary tones were chosen to span a range above and below the resonant frequency of the guinea pig middle ear around 3 to 4 kHz (Mundie, 1963; Relkin, 1988).

Initially, the frequency response in the ear canal to clicks was measured to check if the microphone-earphone assembly tip was correctly fit. If the response did not show a flat curve, the assembly tip was inspected and adjusted until a proper fit was achieved. Then, to calibrate the two earphones, the frequency response to a click from each of the two insert earphones was separately measured and stored to control stimulation level during subsequent recordings. The calibration was repeated each time the pressure in the middle ear was changed. For each middle ear pressure, primary tones of five frequency pairs with the level of the lower-frequency primary (L_1) ranging from 35 to 85 dB SPL ($L_2 = L_1 - 10$) were presented in 5-dB steps. Responses to 100 presentations were averaged for both distortion product otoacoustic emission (DPOAE) or cochlear microphonic distortion product (CMDP) measures.

After data collection, the animal was sacrificed by intraperitoneal injection (ip) of pentobarbital while under general anesthesia. Recordings of DPOAEs and CMDPs were made after cessation of breathing under the same experimental conditions to confirm that data were the result of an active cochlear process rather than the distortion inherent to the experimental system.

Data analysis was based on a model illustrated by the block diagram in Fig. 1(A). The total forward gain $G_f(f,P,Z_{cc})$ is a function of frequency f and middle ear static pressure P, but also depends on middle ear impedance including net cochlear impedance Z_{cc}. The cochlear distortion product generator is represented by the nonlinear device in the middle box. The total backward gain $G_b(f,P,Z_{ec})$ is a function of frequency and middle ear static pressure and depends on middle ear impedance including net ear-canal impedance Z_{ec}. The forward gain in the model is then well-defined in that it is determined by the characteristics of the middle ear and cochlea. The backward gain in this model, however, can be dependent on the canal configuration (Matthews, 1983). In these experiments, we made no attempt to vary the ear canal volume. Rather, we attempted to seal the tip into the canal at a point approximately 4 mm from the tympanic membrane consistently across subjects. The measured levels of distortion products are dependent on this configuration of the ear canal. The results, however, focus on the change in the transmission properties with variations of middle ear pressure, rather than the absolute frequency response. These changes in pressure are likely to change the mechanical properties of the ossicular chain and the tympanic membrane, which would affect signal transmission through the middle ear. To the extent that there are interactions with the external ear, the changes in transmission may also be dependent on the specific configuration of the sealed ear canal.

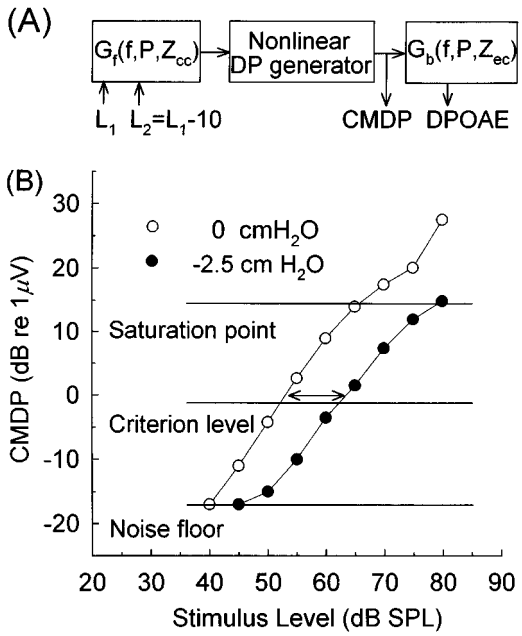


FIG. 1. Signal flow block diagram and forward attenuation in CMDP growth functions. (A) The total forward gain $G_f(f,P,Z_{cc})$ depends on middle ear impedance including net cochlear impedance Z_{cc}, as function of frequency f and middle ear static pressure P. The cochlear distortion product generator is represented by the nonlinear device in the middle box. The total backward gain $G_b(f,P,Z_{ec})$ depends on middle ear impedance including net ear canal impedance Z_{ec}, as a function of frequency and pressure. (B) Attenuation was determined by horizontal distance between the normal (0-cm H_2O) and the abnormal (-2-cm H_2O) middle ear pressure conditions. The effect of pressure change on forward transmission was determined by measuring the horizontal distance between the curve for normal middle ear pressure and curve for abnormal middle ear pressure in the growth function of cochlear microphonic distortion products (CMDPs). The horizontal distance was determined at criterion level of halfway between the noise level and saturation amplitude. Saturation level was determined as the amplitude at which the growth function slope decreases, usually occurring at a stimulus level near 60 dB SPL for normal middle ear pressure.

The model in Fig. 1(A) assumes linear forward and backward transmission properties of the middle ear with a nonlinear distortion component generator in the cochlea. Given this assumed linear transmission, the pressure changes will attenuate forward-transmitted signals, producing a horizontal shift in CMDP growth function. The horizontal distance between the two curves is then a quantitative measure of loss in the forward transmission due to middle ear pressure [Fig. 1(B)]. During backward transmission, if the emission is attenuated by 5 dB due to middle ear pressure, the DPOAE will decrease or drop vertically by 5 dB. Consequently, the DPOAE is affected by both forward and backward attenuation [Fig. 2(A)].

Forward and backward transmission effects were separated out in several steps. First, the amount of forward attenuation was obtained by measuring the horizontal differ-

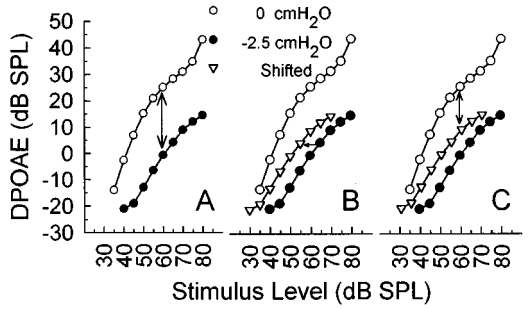


FIG. 2. Measurement of vertical shift for DPOAE growth functions. (A) DPOAE growth functions for two pressure conditions. (B) Shift of the DPOAE growth functions for abnormal middle ear pressure conditions to compensate for forward transmission based on CMDP data. (C) Measurement of vertical shift for horizontally shifted DPOAE growth functions. DPOAE curve for the abnormal pressure condition was horizontally shifted by the amount determined by the CMDP data. The vertical distance between the normal (0-cm H_2O) and the abnormal (-2-cm H_2O) middle ear pressure conditions for shifted functions determined the backward transmission.

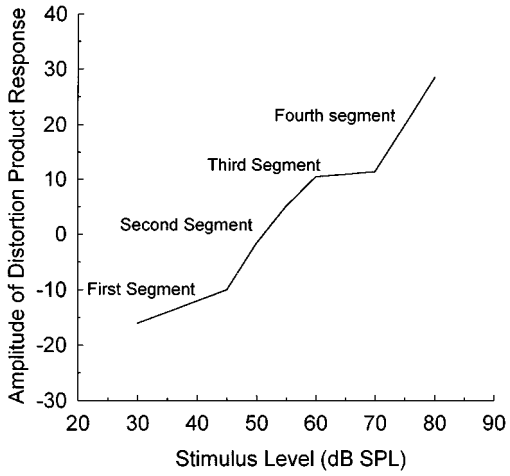


FIG. 3. A schematic growth function with stimulus levels shows a four-segment curve: near noise floor, above noise floor, at saturation, and in steep slope.

ence between the CMDP curves [Fig. 1(B)]. Second, the effect on forward transmission on DPOAE was removed by horizontally shifting the DPOAE curve [Fig. 2(B)] based on the CMDP data [Fig. 1(B)]. Third, the vertical distance was measured between the curve for normal middle ear pressure and the shifted curve for abnormal middle ear pressure [Fig. 2(C)]. This vertical distance then represents the effect of pressure change on backward transmission. The vertical shift for each condition was obtained as an average of vertical distances at stimulus levels of 45, 50, 55, and 60 dB SPL.

The data shown in Figs. 1 and 2 do not conform completely to the assumptions of linearity. The curves in 1(B) are not a simple parallel horizontal shift with changes in pressure and the curves in 2(C) are not a simple vertical shift once corrections for forward transmission are made. These results are representative of the data recorded from other subjects in that the effect on forward transmission generally appears closer to model prediction than that for backward transmission. The results suggest that there may be interactions between the nonlinear and linear elements of the model. One reason for the change in slope may be the effect of middle ear pressure on the inner ear (Maier *et al.*, 1993; Nishihara *et al.*, 1992). A similar phenomenon was found in previous studies, where there was a difference in slope between normal and ± 200-mm H_2O pressures in the ear canal (Naeve *et al.*, 1992) and between normal and diseased middle ear (Owens *et al.*, 1992).

Despite these variations, we viewed the linear approximations as adequate to describe the transmission properties. The variation necessitated specific procedures to define transmission characteristics. A typical growth function can be approximated by four segments as illustrated schematically in Fig. 3. To measure attenuation, we chose a criterion level halfway between the noise floor (first segment) and saturation (third segment) for the following reasons. First, the resulting attenuation values, while not applicable to all stimulus levels, may be indicative of an average attenuation value. Second, the growth in the second segment is most sensitive to metabolic processes and is generally assumed to be produced by an active process in the cochlea (Kim, 1986; Norton, 1992).

The data from each guinea pig were treated as independent samples pooled into one of five abnormal pressure conditions: -10, -5, -2.5, 10, and 20-cm H_2O. The means from each abnormal pressure condition were compared with that from the normal pressure condition (0-cm H_2O) using a statistical test of repeated measures (Winer, 1991). Differences significant at $p<0.05$ are reported.

II. RESULTS

The effect of middle ear pressure change on forward transmission was determined by the CMDP measures as illustrated in Fig. 1(B). Figure 4(A) and (B) shows the average attenuation resulting from middle ear pressure on forward transmission through the middle ear for all subjects. The results in Fig. 4(A) show data for positive pressures as a function of f_2 frequency. The results in Fig. 4(B) show similar data for negative middle ear pressures. Positive pressure sub-

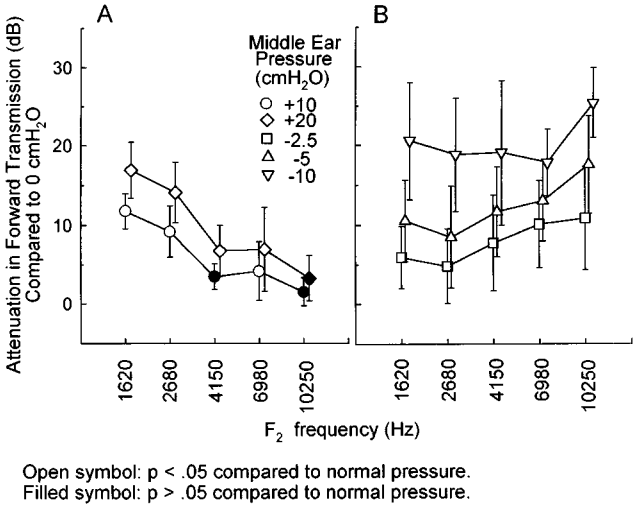


FIG. 4. Average attenuation ($n=20$) in forward transmission as a function of frequency. Parameter is middle ear pressure. The values were determined by the horizontal shift for CMDP growth functions (error bar: ± 1 s.d.).

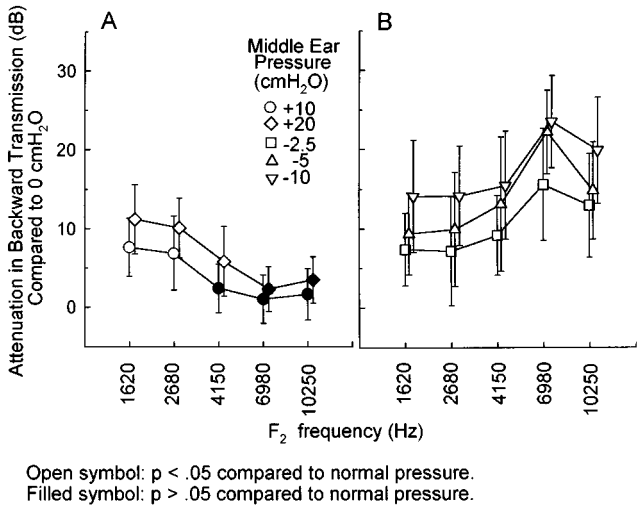


FIG. 5. Average attenuation ($n=20$) in backward transmission as a function of frequency. Parameter is middle ear pressure. The values were determined by the vertical change in the shifted DPOAE growth functions (error bar: ± 1 s.d.).

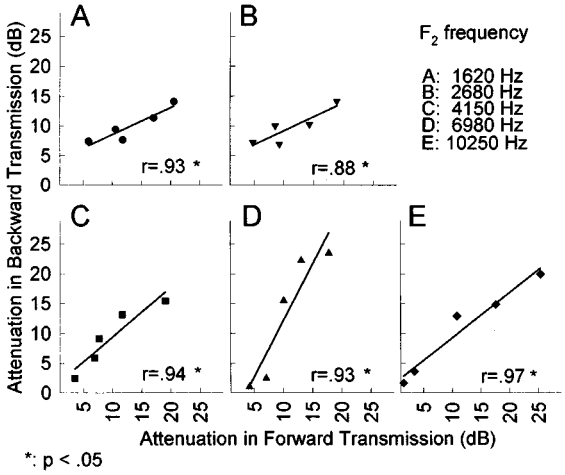


FIG. 6. Scatter plots of forward versus backward attenuation for different frequency stimuli. Calculated correlation coefficients (r) are indicated in each graph. The data are replotted from Figs. 4 and 5.

stantially affected low-frequency transmission, but had less effect on high-frequency transmission. Negative pressure affected both low- and high-frequency transmission. Increasing the magnitude of either the positive or negative pressure increment produced more attenuation.

The effect of middle ear pressure change on backward transmission was determined by the DPOAE measures as illustrated in Fig. 2. Figure 5(A) and (B) shows the effects of abnormal middle ear pressure on backward transmission for all subjects. The results in Fig. 5(A) show the vertical shifts between the normal and positive pressures as a function of f_2 frequency. The results in Fig. 5(B) show the vertical shifts for negative middle ear pressures. Similar to the effects on forward transmission, positive pressure had a greater effect on backward transmission at low frequencies than at high frequencies. Negative pressure had an effect on transmission for both low and high frequencies. Increasing the magnitude of either the positive or negative pressure change resulted in more attenuation. Nevertheless, the effect of negative pressure was generally greater than that of an equivalent positive pressure, especially at high frequencies.

To directly compare forward and backward transmission, attenuation in backward transmission is plotted as a function of the attenuation in forward transmission for the same conditions (Fig. 6). The effects in both measures increase with changes in middle ear pressure. A linear regression fit to the data shows a steeper slope for high frequencies than for low frequencies.

III. DISCUSSION

A. Manipulation of experimental pressure

In previous work, pressure has been manipulated at different sites including the middle ear (Møller, 1965), external ear canal (Robinson and Haughton, 1991), and ambient air within a pressure chamber (Osterhammel *et al.*, 1993). Each may have different effects on emission measurements. Pressure established within a chamber might produce unstable pressure differences across the tympanic membrane because of possible shunting by the eustachian tube function. The pressure in the external canal may have a different effect than pressure in the middle ear in terms of the position of tympanic membrane and ossicles. Positive pressure in the external canal or negative pressure in the middle ear will have the same effect on the eardrum. But, the former may not induce a substantial negative pressure in the middle ear. Negative middle ear pressure can cause disarticulation of ossicular chain (Mundie, 1963; Møller, 1965). However, it has not been established whether or not the oval window and round window membrane are pulled toward the middle ear.

B. Effects of positive middle ear pressure

Forward and backward transmission were similarly affected by positive pressures in the middle ear (Figs. 4 and 5). Low frequencies ($f_2=1620$ and 2680 Hz) were more affected by positive middle ear pressure than were high frequencies ($f_2=6980$ and 10 250 Hz). These results are in agreement with other studies which showed that low frequencies were affected by pressure changes (Kemp *et al.*, 1990; Robinson and Haughton, 1991; Hauser *et al.*, 1991, 1992; Naeve *et al.*, 1992; Osterhammel *et al.*, 1993).

When the pressure becomes positive in the middle ear, the tympanic membrane is pushed toward the external ear and its stiffness is increased (Møller, 1965). Since the middle ear is mainly dominated by stiffness at low frequencies, it is expected that low frequencies would be more affected than high frequencies. This effect is well documented for forward transmission (Møller, 1965). Increasing positive pressure (from 10 to 20-cm H_2O) further decreased transmission through the middle ear. Other studies have also shown that the effect on OAE amplitudes increased with increasing magnitude of middle ear pressure (Robinson and Haughton, 1991; Naeve *et al.*, 1992; Osterhammel *et al.*, 1993).

C. Effects of negative middle ear pressure

Negative middle ear pressures had a greater effect on both forward and backward transmission than positive pres-

sures (Figs. 4 and 5). Similar results have also been observed in previous studies. Møller (1965) measured the cochlear microphonic in cats and observed that negative pressure had a greater effect on signal attenuation for almost all frequencies from 250 to 6000 Hz. In a study of external pressure effects on human DPOAE measures, Osterhammel *et al.* (1993) showed that negative pressure in the external ear canal produced a greater effect for frequencies of 2, 6, and 8 kHz Owens *et al.* (1992) measured both transiently evoked OAEs (TEOAEs) and DPOAEs for one subject who suffered from abnormal middle ear pressure. The measurements were made when the middle ear pressure was at approximately 100 daPa (about 10-cm H_2O) the first time and at about -100 daPa five weeks later. TEOAEs for both positive and negative pressures were not measurable. The amplitudes of DPOAEs were much smaller for 2000 Hz (geometric mean of primary frequencies) at -100 daPa than at 100 daPa.

D. Effects on high frequencies

High-frequency transmission was also affected by negative pressures (Figs. 4 and 5). Møller (1965) showed that the pressure change also affected frequencies between 4 and 6 kHz. Osterhammel *et al.* (1993) showed that the amplitude of emission at 8 kHz decreased by approximately 10 dB, more than that at 2–6 kHz. The data obtained by Naeve *et al.* (1992) for TEOAE measures showed that emission energy began to decrease in the 4- to 5-kHz region. No data were shown for frequencies above 5 kHz due to the filtering used in TEOAE measures. Magnan and Vassout (1992) measured cochlear microphonic (CM) with positive pressure in the guinea pig ear canal. They found that the maximum effect was between 8 and 10 kHz (-25 dB) with less effect at low frequencies. Lonsbury-Martin *et al.* (1993) showed an example of a subject with negative pressure where the DPOAE amplitude decreased up to 8000 Hz. They reported that it was common under acute negative pressure conditions, but they did note one case where only low-frequency OAEs were affected by negative middle ear pressure.

The mechanism underlying the effect of negative pressure appears to be more complicated for high-frequency transmission. Based on a simple mechanical model of the middle ear, increasing the stiffness caused by negative pressure should not strongly affect high-frequency transmission which is mainly mass-controlled. However, there are several other ways in which high-frequency transmission might be affected. For instance, the ossicular chain can flex at high frequencies, reducing transmission (Guinan and Peake, 1967). Such flexure may increase due to changes in the position of the ossicular chain. There is also a possibility of decoupling the stapes from the incus at high frequencies due to the pressure change (Nuttall and Ross, 1990), especially with negative pressure (Mundie, 1963; Møller, 1965). Møller (1965) observed differences in the effect of negative from positive pressure on acoustic resistance. When the ossicular chain is decoupled, acoustic resistance decreases, but the attenuation in transmission through ossicular chain increases. Another possibility is that negative pressure compresses the ossicular chain or results in contact with other structures, such as the wall of the middle ear cavity or the oval window.

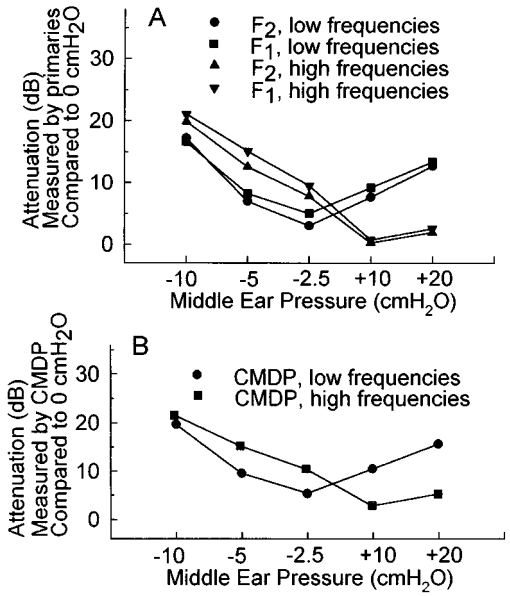


FIG. 7. The effects of middle ear pressure on forward transmission using three different measures are compared. The same stimulus is used in each cases, but middle ear attenuation is assessed by one of the two primary frequencies (f_1 or f_2) or by the CMDP. In each case the effect is evaluated for high-frequency stimuli (where $f_2 = 6980$ and 10 250 Hz) and for low-frequency stimuli (where $f_2 = 1620$ and 2680 Hz). The ratio of f_2/f_1 is fixed at 1.2. Part (A) shows the attenuation measured using the primary frequencies, labeled either f_1 or f_2. Part (B) shows the attenuation measured using the distortion product frequency (CMDP).

Such changes in the ossicular chain could effectively result from movement of both tympanic membrane and footplate toward the middle ear cavity by negative pressure. A mechanical contact could increase the mass of the ossicles resulting in an effect on high-frequency transmission.

E. CM primaries

In this study, we used CMDP to assess the effect on forward transmission. The measurement of CM primaries can be an alternative to assess the same effects. Comparisons of the results using these two methods of assessing middle ear transmission changes are shown in Fig. 7. Figure 7(A) illustrates the measures for both primary tones. Figure 7(B) shows similar measures for the CMDP. The similarity among these three measures is evident, suggesting that the particular method used to evaluate transmission is not a major factor in the results.

The change in transmission of the primaries can have an influence on the effective levels of f_1 and f_2 reaching the cochlea. The ratio of L_1/L_2 was set at 10 dB because our preliminary experiments showed that at that ratio, the distortion product $2f_1 - f_2$ was robust. Although we set the ratio at 10 dB for the acoustic signal under all conditions, the ratio of the recorded CM at f_1 and f_2 varied. Figure 7 illustrates the attenuation in CM produced by middle ear pressure at both low and high primary frequencies. Since the two primaries were close in frequency, the change in the CM primary ratio for both positive and negative pressure conditions was limited to a range of approximately 2 dB, i.e., the ratio of CM primaries was in the range of 8–10 dB across condi-

tions. These changes in attenuation could potentially have an impact on the recorded distortion product, however, the optimum difference L_1/L_2 that produces maximum DPOAE amplitude is relatively unaffected by the middle ear pressure (Richter *et al.*, 1994).

F. Clinical applications

The study confirms that abnormal pressure in the middle ear attenuates OAE amplitudes. The greater the abnormal pressure, the greater the effect. A low amplitude of OAE could be the result of a mild abnormal middle ear pressure or a mild cochlear disorder. The measurement of OAEs could be used to monitor the severity of middle ear disorders and the result of the treatment. This study further showed that negative pressure had a significant effect on transmission at frequencies above 6 kHz. An unmeasurable OAE at high frequencies, therefore, does not necessarily mean that the cochlea is abnormal. Some positive pressure significantly affects transmission at low frequencies, the midfrequency (4 to 5 kHz) may be least affected by middle ear pressure changes.

The effects of middle ear pressure on forward and backward transmission were similar (Fig. 6). This may result in a smaller change in behavioral hearing threshold than that in OAE threshold because behavioral hearing threshold depends only on forward attenuation. We reason that measures such as ABR and behavioral testing may be affected less by middle ear pressure and therefore may be more suitable in cases of middle ear disease. When OAEs are not measurable due to an abnormal middle ear pressure, the stimulus level could be increased to some extent to compensate. However, the decreased growth slope with pressure change in OAEs suggests that the utility of increasing stimulus level may be limited.

G. Summary

This study demonstrates a method to separately evaluate the effects of middle ear pressure on forward and backward transmission in an animal model. The effects on forward and backward transmission were similar. Positive pressures affected transmission for low-frequency stimuli but had little effect for high-frequency stimuli. Negative pressures had a greater effect at low frequencies and significantly affected high-frequency transmission. Reduction of OAE is dependent on the severity of abnormal middle ear pressures.

ACKNOWLEDGMENTS

The authors wish to thank S. Neely and Z. Liu at Boys Town National Research Hospital for their software (EMAV v3.1) and G. Haskell at Iowa City V.A. Hospital for use of recording equipment. We would also like to acknowledge the comments of W. Peake and R. Margolis who were reviewers for an earlier version of this manuscript.

Fria, T. J., Gantekin, E. I., and Eichler, J. A. (**1985**). ''Hearing activity of children with otitis media with effusion,'' Arch. Otolaryngol. **111,** 1–16.

Guinan, J. J., and Peake, W. T. (**1967**). ''Middle ear characteristics of anesthetized cat,'' J. Acoust. Soc. Am. **41,** 1237–1261.

Harvey, R. C., and Walberg, J. (**1987**). ''Special considerations for anesthesia and analgesia in research animals,'' in *Principles and Practice of Veterinary Anesthesia*, edited by C. E. Short (Williams & Wilkins, Baltimore).

Hauser, R., Probst, R., and Harris, F. P. (**1991**). ''Effects of variation in atmospheric pressure on the amplitude of distortion-product otoacoustic emissions in humans,'' *Abstracts of the 14th Midwinter Research Meeting* (Association for Research in Otolaryngology, St. Petersburg, FL), p. 66.

Hauser, R., Probst, R., Harris, F., and Frei, F. (**1992**). ''Influence of general anesthesia on transiently evoked otoacoustic emissions in humans,'' Ann. Otol. Rhinol. Laryngol. **101,** 994–999.

Howie, V. M. (**1975**). ''Natural history of otitis media,'' Ann. Otol. Rhinol. Laryngol. Suppl. **19,** 67–72.

Kemp, D. T., Ryan S., and Bray, P. (**1990**). ''A guide to the effective use of otoacoustic emissions,'' Ear Hear. **11**, 93–104.

Kim, D. O. (**1986**). ''Active and nonlinear cochlear biomechanics and role of outer-hair-cell subsystem in the mammalian auditory system,'' Hearing Res. **22,** 105–114.

Lonsbury-Martin, B. L., McCoy, M. J., Whitehead, M. L., and Martin, G. K. (**1993**). ''Clinical testing of distortion product otoacoustic emissions,'' Ear Hear. **1**, 11–22.

Magnan, P., and Vassout, P. (**1992**). ''Modification of the middle ear transfer function by static pressure applied on the tympanic membrane,'' *29th Workshop on Inner Ear Biology*, Engelberg, Switzerland (unpublished), p. 95.

Maier, W., Beck, C., and Hauser, R. (**1993**). ''Effects of a change in environmental pressure on spontaneous nystagmus in patients with cochleovestibular diseases,'' Laryngo-Rhino-Otologie. **72**, 236–241.

Matthews, J. W. (**1983**). ''Modeling reverse middle ear transmission of acoustic distortion signals,'' in *Mechanics of Hearing*, edited by E. de Boer and M. A. Viergever (Delft U. P., Delft), pp. 11–18.

Møller, A. R. (**1965**). ''An experimental study of the acoustic impedance of the middle ear and its transmission properties,'' Acta Oto-laryngol. **60,** 129–149.

Mundie, J. R. (**1963**). ''The impedance of the ear—a variable quantity,'' in *Middle Ear Function Seminar*, edited by M. J. L. Fletcher, U.S. Army Medical Research Laboratory Report No. 576, pp. 63–85.

Naeve, S. L., Margolis, R. H., Levine, S. C., and Fournier, E. M. (**1992**). ''Effect of ear-canal air pressure on evoked otoacoustic emissions,'' J. Acoust. Soc. Am. **91,** 2091–2095.

Nishihara, S., Gyo, K., and Yanagihara, N. (**1992**). ''Transmission of change in the atmospheric pressure of the external ear to the perilymph,'' Am. J. Otol. **13**, 364–368.

Nuttall, A. L., and Ross, M. D. (**1990**), ''Auditory physiology,'' in *Otolaryngology*, edited by G. M. English (Harper & Row, Philadelphia), pp. 18, 19.

Norton, S. J. (**1992**). ''Cochlear function and otoacoustic emissions,'' Seminar in Hearing **13**, 1–14.

Norton, S. J. (**1993**). ''Application of transient evoked otoacoustic emissions to pediatric populations,'' Ear Hear. **14**, 64–73.

Osterhammel, P. A., Nielsen, L. H., and Rasmussen, A. N. (**1993**). ''Distortion product otoacoustic emissions,'' Scand. Audiol. **22,** 111–116.

Owens, J. J., McCoy, M. J., Lonsbury-Martin, B. L., and Martin, G. K. (**1992**). ''Influence of otitis media of evoked otoacoustic emissions in children,'' Seminar in Hearing **13**, 53–65.

Relkin, E. M. (**1988**). ''Introduction to the analysis of middle-ear function,'' in *Physiology of the Ear*, edited by F. Jahn and J. Santos-Sacchi (Raven, New York), pp. 103–122.

Richter, B., Hauser, R., and Lohle, E. (**1994**). ''Dependence of distortion product emission amplitude on primary-tone stimulus levels during middle-ear pressure changes,'' Acta Oto-Laryngol. **114,** 278–284.

Robinson, P. M., and Haughton, P. M. (**1991**). ''Modification of evoked oto-acoustic emissions by changes in pressure in the external ear,'' Br. J. Audiol. **25**, 131–133.

Teele, D. W., Klein, J. O., and Rosner, B. A. (**1980**). ''Epidemiology of otitis media in children,'' Ann. Otol. Rhinol. Laryngol. **89,** 5–6.

Trine, M. B., Hirsch, J. E., and Margolis, R. H. (**1993**). ''The effect of middle ear pressure on transient evoked otoacoustic emissions,'' Ear Hearing **14**, 401–407.

Winer, B. J. (**1991**). *Statistical Principles in Experimental Design* (McGraw-Hill, New York).

Intermodulation components in inner hair cell and organ of Corti responses

M. A. Cheatham[a] and P. Dallos
Hugh Knowles Center, Department of Communication Sciences and Disorders, Frances Searle Building, Northwestern University, Evanston, Illinois 60208-3550

(Received 23 December 1996; revised 2 April 1997; accepted 3 April 1997)

Two-tone responses are recorded from inner hair cells and from the organ of Corti fluid space in second and third turns of the guinea pig cochlea where best frequencies (BF) are approximately 4000 and 1000 Hz, respectively. This allows both ac and dc response components to be obtained and facilitates comparisons with psychophysical investigations that have traditionally been conducted at low and moderate frequencies. The measurements of ac responses in the organ of Corti fluid space also allow comparisons with mechanical results because the cochlear microphonic is proportional to basilar membrane displacement. By using a constant frequency ratio (f_2/f_1) of 1.4, *local* distortion products generated at the recording location are prominent when the two primaries are near the BF of the cell. However, when the primary pairs increase above BF, quadratic and cubic difference tones are recorded even when responses to the primaries are not measurable. The presence of these *traveling* distortion products is consistent with the idea that both f_2-f_1 and $2f_1-f_2$ have their own traveling waves. Notches in the existence regions of quadratic and cubic difference tones were also observed and found to be influenced by mutual suppression between the two inputs. © *1997 Acoustical Society of America.* [S0001-4966(97)01908-5]

PACS numbers: 43.64.Ld, 43.64.Nf, 43.64.Tk [RDF]

INTRODUCTION

Approximately 150 years ago, Ohm (1843) proposed that combination tones are perceived by resolving a complex sound into its individual Fourier components. He believed that an individual pitch could be distinguished only when the stimulus contained the corresponding frequency. Helmholtz (1954) later suggested in 1863 that a mechanical nonlinearity in the middle ear introduces distortion products which are then perceived behaviorally as if they had been presented at the input to the ear. Evidence for Ohm's original notion was presented by Zwicker (1955) who demonstrated that a pitch at the cubic difference tone (CDT) frequency, $2f_1-f_2$, could be made to disappear by adjusting the amplitude and phase of an external tone whose frequency matched that of the CDT. A physiological correlate for these psychophysical results was provided in 1968 when Goldstein and Kiang (1968) reported time-locked responses to the CDT in single auditory nerve fibers recorded in the cat. It was also possible to cancel this response to the intermodulation component by adding a third tone to the stimulus which was equal in frequency to the distortion product. Goldstein and Kiang also demonstrated that even if f_1 and f_2 were outside the response area of the fiber, the response to the CDT was robust when the latter was near the best frequency (BF) of the cell. The implication of these results is that the difference tone is generated in the region of overlap between the two primaries. Energy at the distortion frequency is then fed back into the mechanics and resolved at its proper place along the cochlear partition.

These ideas were supported by Kim *et al.* (1980) who studied the spatial distribution of cochlear responses to two-tone stimuli by recording from hundreds of nerve fibers in a single cat. At low levels, the difference tones, f_2-f_1 and $2f_1-f_2$, emerge in the primary frequency region. With increasing level, these components show additional activity at their characteristic place so that they are clearly visible in the distortion frequency region as well as in the frequency region corresponding to the primaries, f_1 and f_2.

Although the exact site of the nonlinearity is not known, it certainly does not reside in the middle ear (Lewis and Reger, 1993; Wever *et al.*, 1940; Guinan and Peake, 1967) as Helmholtz originally surmised. This is because the magnitude of these components is strongly dependent upon the frequency separation between the two primaries (Goldstein, 1967; Goldstein and Kiang, 1968). In addition, it has been demonstrated (Smoorenburg, 1972) that the CDT is perceived only when both primaries are audible. Consequently, any defect which causes a threshold shift is thought to precede the nonlinearity. In other words, stimulus components below the level of audibility do not reach the distortion generator. Smoorenburg's results, obtained in a patient with threshold shift, were subsequently confirmed in a behaviorally trained chinchilla (Dallos, 1977) with an outer hair cell (OHC) loss produced by the ototoxic antibiotic, kanamycin. In fact, outer hair cell damage appears to linearize cochlear output (Dallos *et al.*, 1980; Patuzzi *et al.*, 1989) implying that the nonlinearity occurs prior to inner hair cell (IHC) transduction.

Based on these ideas, it is of interest to learn whether local and traveling distortion products can be resolved in IHC receptor potentials. For example, when the ear is stimulated with two primaries whose intertone $[(f_1+f_2)/2]$ corre-

[a]Corresponding author address: 2-240 Frances Searle Building, 2299 North Campus Drive, Northwestern University, Evanston, IL 60208-3550. Electronic mail: m-cheatham@nwu.edu

sponds to the BF of the cell, a profusion of combination tones is measured. These responses reflect local components produced at the generation site. However, if the primaries are presented well above BF, such that the difference tone frequency coincides with the BF of the IHC, then traveling components should be resolved. The latter are thought to originate basal to the recording location and, via their own traveling waves, propagate apically to stimulate hair cells whose BFs correspond to the individual distortion product frequencies (Goldstein, 1967). Although traveling components have been demonstrated in the cochlear microphonic (CM) (Gibian and Kim, 1979, 1982; Dallos *et al.*, 1980) and in the dc receptor potentials of IHCs with high BFs (Nuttall and Dolan, 1990), no data are available for IHCs with lower BFs or for the ac receptor potential. The latter, as well as extracellular measures of both cubic and quadratic components, are important when trying to compare hair cell and mechanical responses (Nuttall *et al.*, 1990; Robles *et al.*, 1991, 1997). It is also important to study cochlear nonlinearities in apical regions to foster comparisons with psychophysical investigations that have usually been restricted to low and moderate stimulus frequencies but have been compared to high-frequency physiological data.

I. METHODS

To avoid repetition, we provide a cursory description of the methods employed in these experiments. Further details can be obtained from previous publications (Dallos *et al.*, 1982; Dallos, 1985; Cheatham and Dallos, 1992). As before, all animal care procedures were approved by the National Institutes of Health and by Northwestern University's Institutional Review Committee.

Young albino guinea pigs were anesthetized with urethane or with a combination of sodium pentobarbital and Innovar-Vet. The standard ventro-lateral approach (Dallos *et al.*, 1982) allowed access to the right auditory bulla which was opened widely. A small window was then made in the cochlear bone over scala media. The use of backlighting assisted electrode placement. After traveling through the endolymphatic space, Hensen's cells are encountered at the peripheral edge of the organ of Corti. Further advances allow recordings from individual hair cells as well as from the organ of Corti fluid space. Although data were collected in both second and third turns, only one opening was made in any given preparation.

The two-tone inputs used in these experiments were generated by a single dynamic earphone (Beyer DT-48) with the result that the two signals were electrically and not acoustically mixed. Sound-pressure levels in the external ear canal were determined using a subminiature microphone (Knowles BT-1751). The latter was attached to a probe tube which was inserted into the sound tube in a concentric arrangement. Harmonic and intermodulation distortion was measured in a coupler using a 1/2 in. B&K condenser microphone and found to be at least 60 dB down from the two stimulating primaries. These determinations, however, were made at maximum sound-pressure levels which were above those used to collect the data reported here. Consequently, these estimates are thought to be conservative.

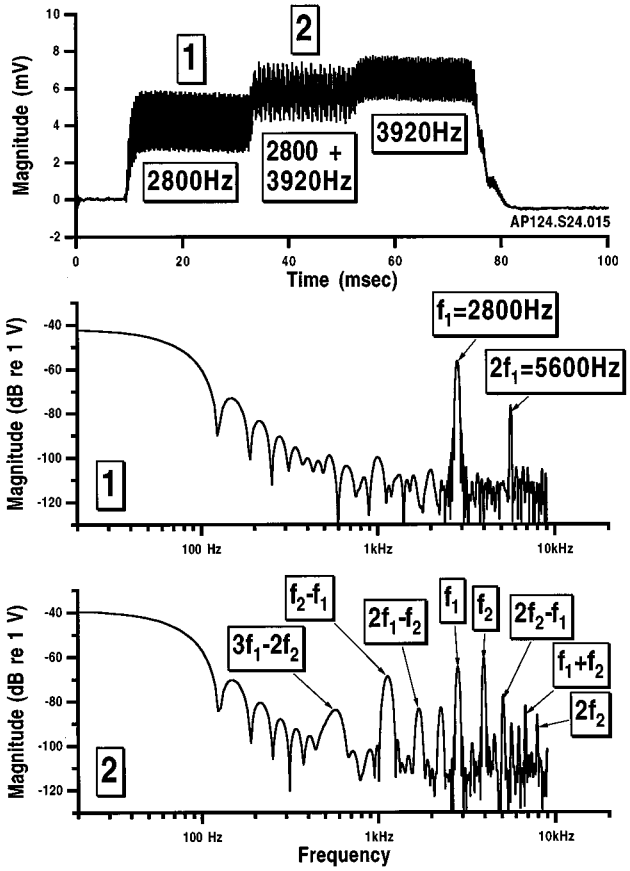


FIG. 1. The averaged response waveform at the top was recorded in a second turn IHC at a sound-pressure level of 70 dB. This and all sound-pressure levels are measured *re*: 20 μPa. The timing of the two inputs shown here is arranged so that responses to each individual primary can be recovered as well as responses to various components in the region of overlap. The center panel shows the spectrum obtained in window 1 for f_1 alone at 2800 Hz. When the FFT is obtained from window 2, the spectrum contains several distortion products with frequencies both above and below the primaries.

All responses were amplified and capacitance compensated (Dagan 8700), low-pass filtered to prevent aliasing and averaged for off-line analysis using Igor Pro (WaveMetrics, Lake Oswego, OR 97035) and/or custom-designed software. An automatic gain control system was also used to optimize amplification, thereby avoiding saturation of the analog-to-digital converter. The peak values of various response components were determined from fast Fourier transforms (FFT) of averaged response waveforms. Waveform segments of approximately 20 ms duration were windowed before transformation using a Hanning function. In some instances, peak ac values were increased by +6 dB/octave above a corner frequency of 470 Hz (Dallos, 1984) to compensate for filtering by the cell's basolateral membrane and by +12 dB/octave above 3500 Hz to compensate for losses in the recording apparatus (Baden-Kristensen and Weiss, 1983; Cody and Russell, 1987). Although these adjustments have been employed before (Cheatham and Dallos, 1993), they yield qualitative, not quantitative, information.

The protocol used in these experiments is provided in Fig. 1. This recording was made in second turn where BFs are approximately 4000 Hz. Because the onset of the second

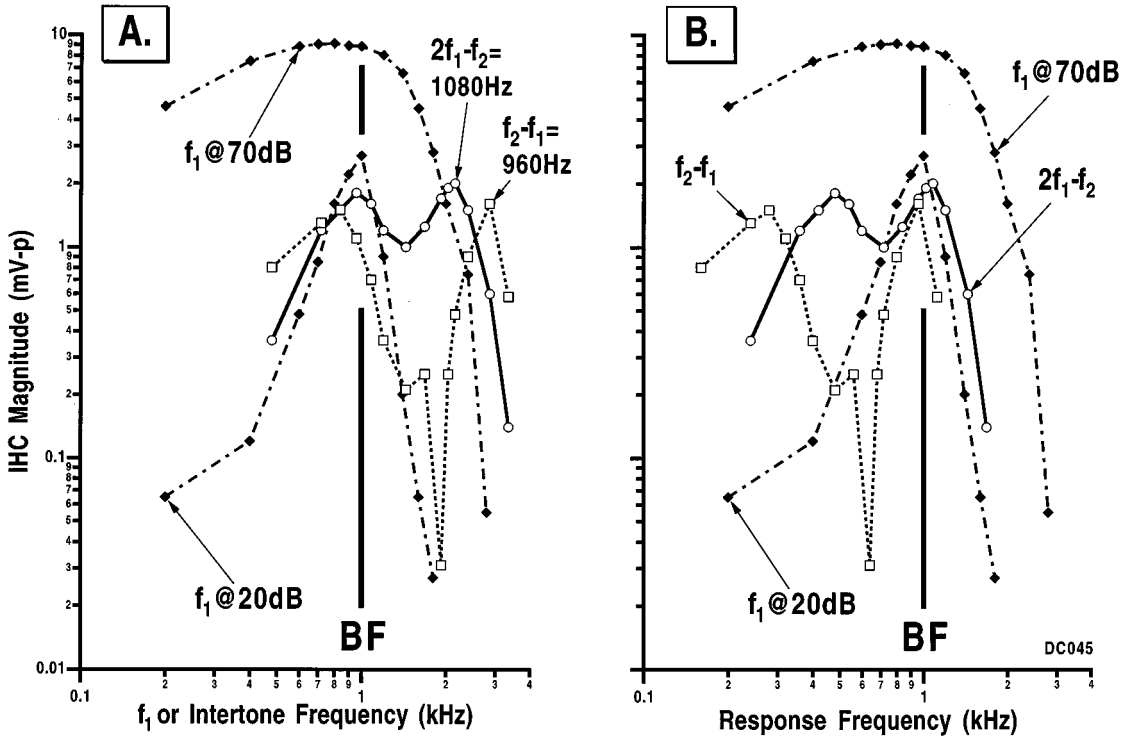


FIG. 2. Frequency response functions at 70 dB are provided for a third-turn IHC. The broken vertical line at 1000 Hz indicates the BF of the cell. In both panels (A) and (B), peak ac values (mV-*p*) are plotted for the lower frequency primary alone (dot-dashed lines) along with two-tone functions (open symbols) for components below f_1 at $2f_1-f_2$ and f_2-f_1. In addition to the frequency response function for f_1 alone at 70 dB, a similar function is included for single-tone inputs at 20 dB to facilitate comparisons between single- and two-tone responses. Data in (A) are plotted at the intertone; those in (B) at the frequency of the individual response component.

tone at 3920 Hz is delayed relative to that of the lower frequency tone at 2800 Hz, it is possible to recover responses to each primary alone as well as those in the region of overlap where both tones are presented together. For example, the spectrum of the f_1 alone response in the center shows magnitude peaks for f_1 at 2800 Hz and its second harmonic at 5600 Hz. In contrast to this simple picture, the spectrum of the two-tone response at the bottom reveals a profusion of response peaks for difference and summation tones as well as harmonic components. Above the primaries, the combination tones f_2+f_1 and $2f_2-f_1$ as well as the second harmonic can be identified. In the region below f_1, magnitude peaks appear at f_2-f_1 and $2f_1-f_2$. Based on the power series approximation, these latter components are referred to as the quadratic difference tone (QDT) and cubic difference tone (CDT). The unlabeled peak at 2240 Hz, just below f_1, could reflect one of a number of components produced by interactions between a primary and a combination tone or between two individual distortion products.

In this example, and in all materials, the primaries were presented at equal levels. This choice was based on input–output functions for the cochlear microphonic recorded in third turn where the largest cubic difference tone was produced at or near the point where the intensities of both primaries were equal (Dallos, 1969). In addition, the frequency ratio f_2/f_1 was kept constant at 1.4 except when small adjustments were made to assure that the two inputs were not harmonically related. This relatively large frequency ratio was used so that the primaries were not excitatory when quadratic and cubic difference tones were near the BF of the IHC. If a smaller frequency ratio at 1.2 had been employed, the primaries would need to be very high at 20 000 and 24 000 Hz to generate a QDT at 4000 Hz. In addition, to produce a CDT at 4000 Hz, the primaries at 5000 and 6000 Hz would probably cause excitation at the 4000 Hz place. Consequently, by using a frequency ratio of 1.4, traveling difference tones below f_1 can be recovered without being contaminated by excitation due to the primaries at the distortion product place.

Although this choice of stimulus parameters has advantages, it is not optimal for all components especially those above the f_2 frequency. For example, when summation tones fall within the response area of the cell, the generating primaries also produce sizeable responses when presented alone. For primary pairs near and above BF, the summation tones as well as $2f_2-f_1$ are usually removed by filtering at the generation site. Hence, in the present experiments, it is not possible to distinguish local from traveling contributions. When responses are measured for these components, as in Fig. 1, they appear to reflect those produced locally for intertones near and below BF.

II. RESULTS

A. Traveling versus local difference tones

Two-tone results are introduced using an IHC recording from third turn where BF is ~1000 Hz. The ac frequency response functions in Fig. 2 are plotted for single- and two-

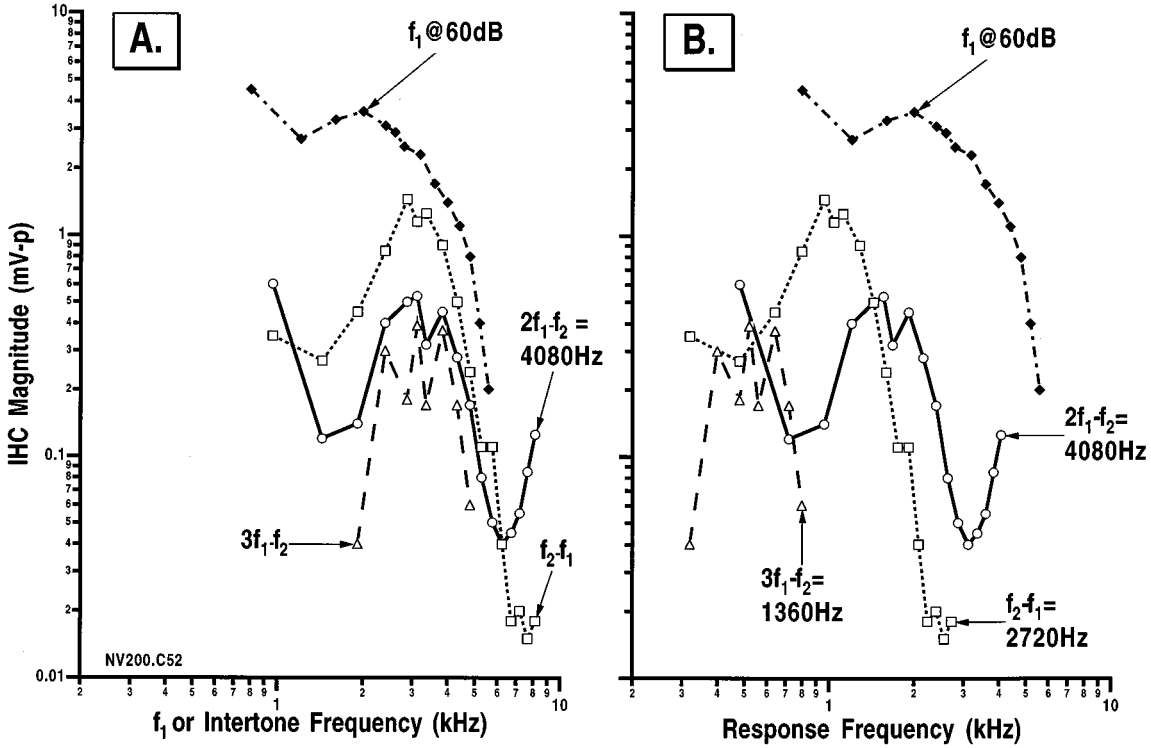


FIG. 3. Results shown here were obtained from a second-turn IHC at 60 dB. The components plotted are as in Fig. 2 except that three difference tones below f_1 are included. A frequency response function for f_1 alone is also provided for reference.

tone inputs at 70 dB. A second, single-tone function is appended to illustrate the response area at 20 dB. In panel (A), on the left, the functions are plotted either at f_1 for the single-tone functions or at the intertone, the average frequency of the two stimulating primaries, for the difference tones at f_2-f_1 and $2f_1-f_2$. The latter were generated by equal-intensity primary pairs that increased in frequency to map out the existence regions for individual distortion products. The two-tone functions, plotted for quadratic and cubic difference tones, exhibit two magnitude peaks. The first is produced by primaries with an intertone at 960 Hz which is near the BF of the cell as indicated by the broken vertical line at 1000 Hz. Here the CDT is 480 Hz. This local component is produced at the generation site. The second peak is associated with a CDT at 1080 Hz and with an intertone at 2040 Hz which is well above BF. A similar behavior is shown for the QDT at 960 Hz where the first peak occurs when $\mathrm{BF}=(f_1+f_2)/2$; the second peak, when $\mathrm{BF}=f_2-f_1$. To distinguish responses produced by primary pairs well above BF from those recorded at the generation site by lower-frequency primaries within the response area of the cell, the former are referred to as *traveling* rather than *local* components.

In order to emphasize that large receptor potentials are produced when the distortion product coincides with the BF of the cell, the results are plotted in panel (B), not at the intertone, but at the frequency of each individual difference tone. Single-tone functions at 70 and 20 dB are again provided for reference. Both quadratic and cubic components exhibit peaks around 1000 Hz, the BF of the cell. It is assumed that these responses are generated at a more basal location and are distributed via their own traveling waves to the more apical location where they are recorded. The tuning and magnitude of these distortion products is commensurate with that observed for single tones at 20 dB (Dallos and Cheatham, 1992, Fig. 11). Comparisons with single-tone inputs, however, are made with caution because distortion products can suffer from the suppression that occurs between multiple components. In this plot (B), both difference tones also have peaks below BF when the intertone frequency is near 1000 Hz. These maxima, corresponding to local components produced at the recording location, peak at BF in plot (A). For these measures at 70 dB, the magnitudes of the difference tones are similar for local and traveling distortion components and for the quadratic and cubic components. The results plotted in Fig. 2 emphasize that local distortion components are generated when the intertone is near the BF of the cell, as in panel (A), while traveling components are expressed when the frequency of the distortion product approximates BF as in panel (B).

Companion measures for a second turn IHC are displayed in Fig. 3. Again, in order to provide an indication of the response area of the cell, the ac frequency response function for f_1 alone at 60 dB is provided. In addition, several difference tones with frequencies below f_1 are also plotted at the intertone in panel (A). These three functions all exhibit peaks when the intertone is near BF indicating the local nature of these responses. In the region above BF, all components decrease as primaries move out of the response area of the cell. However, with further increases, the response for $2f_1-f_2$ begins to rise when the distortion product is around the BF of the cell at 4080 Hz. This occurs when the intertone

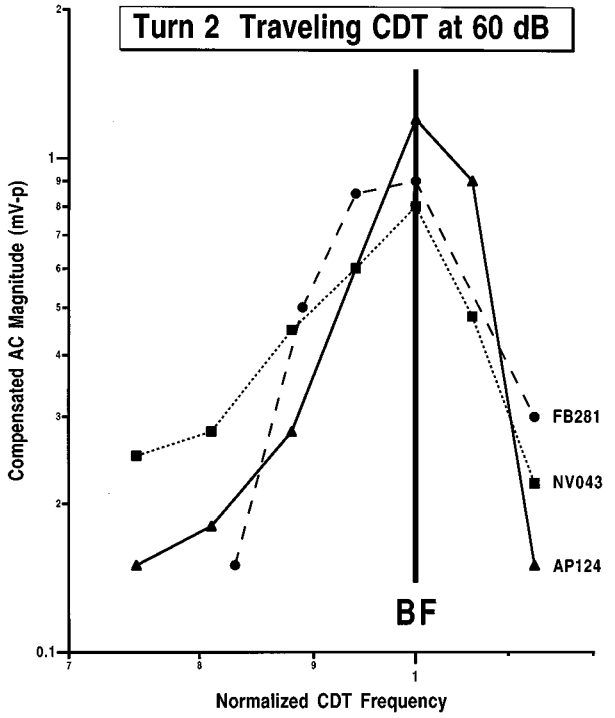


FIG. 4. Data from three IHCs located in second turn are plotted here on a normalized frequency scale where BF is indicated by the vertical bar at 1.0. The ac voltages have been compensated for filtering by the cell's basolateral membrane and by the microelectrode. These responses at 60 dB are plotted for the CDT produced in response to primaries placed well above BF. Consequently, the functions peak near BF when the CDT is near 4000 Hz.

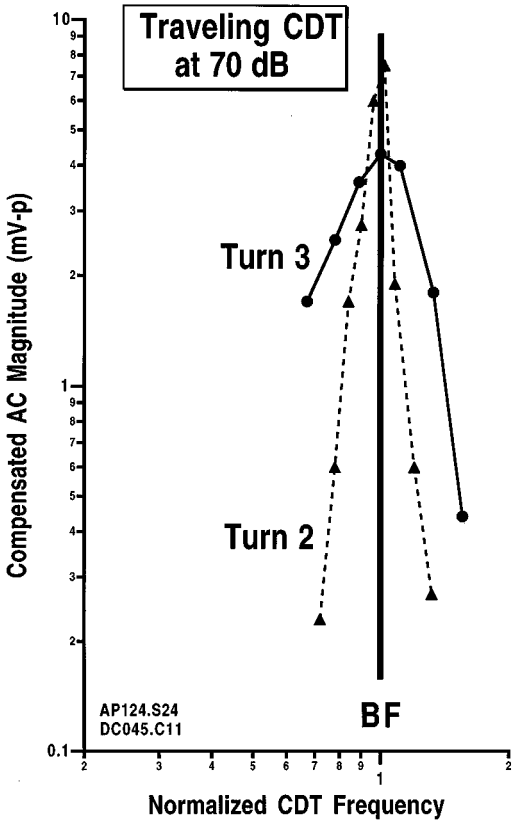


FIG. 5. These results are similar to those in Fig. 4 except that examples are provided here for IHCs in second and third turns. These responses were measured at 70 dB.

frequency is 8160 Hz. These results are also plotted at the frequency of the distortion product in panel (B) to emphasize that components with frequencies at and around the BF of the cell behave very much as single-tone inputs of the same frequency. Unfortunately, in this example, the primary frequencies were not extended high enough to define the traveling CDT or to expose those for the other components. In other words, it is only the CDT frequency that approximates the BF of the cell. Consequently, additional examples are provided to explore the responses generated by higher-frequency, primary pairs.

This information is presented in Fig. 4 for three second-turn IHCs where the primaries are also at 60 dB. Since ac responses are filtered by the resistances and capacitances associated with the hair cell's basolateral membrane, they do not provide a good indication of the relative magnitudes of individual response components especially when recordings are made from more basal regions of the cochlea. Consequently, ac receptor potentials are compensated for these reductions, as well as those associated with the recording apparatus, using the corrections provided in Cheatham and Dallos (1993). Compensated ac responses are also plotted on a normalized frequency scale obtained by dividing the frequency of the difference tone by BF. However, in contrast to the third turn data plotted in Fig. 2, these estimates of BF were not determined from low-level responses in individual cells but are based on an average of cells recorded in this region. Notice that all functions peak near 1.0 when the frequency of the CDT corresponds to the BF of the recording location. Since these components are probably generated at a more basal location, the intertone frequencies associated with the magnitude peaks are well above BF at either 7680 or 8160 Hz. The functions have Q_{10}'s of ~4.0 which is consistent with single units recorded from the guinea pig auditory nerve in this frequency region (Evans, 1972).

The primary motivation behind compensating ac responses was to foster comparisons between IHCs at different cochlear locations as well as between IHC and basilar membrane responses. Consequently, results from turns 2 and 3 are provided in Fig. 5. In order to facilitate these comparisons, the frequency scale has again been normalized by dividing the distortion product frequency by BF which is represented by the vertical line at 1.0. Data are plotted for the traveling CDT with the result that the functions peak at 1.0. This occurs when the primaries are well above BF so that the distortion product approximates the BF of the cell. As one might expect, the data from turn 3 where BF is approximately ~1000 Hz, reflect broader tuning than those from the second turn where BF is ~4000 Hz. In fact, the Q_{10} for the third turn in ~1.3 while that for the second turn is ~3.7. Both of these values fall within the ranges reported for single units in the guinea pig auditory nerve (Evans, 1972). Since these traveling distortion products behave as if a third tone had been added to the stimulus, they should reflect the filter shape exhibited at their BF locations.

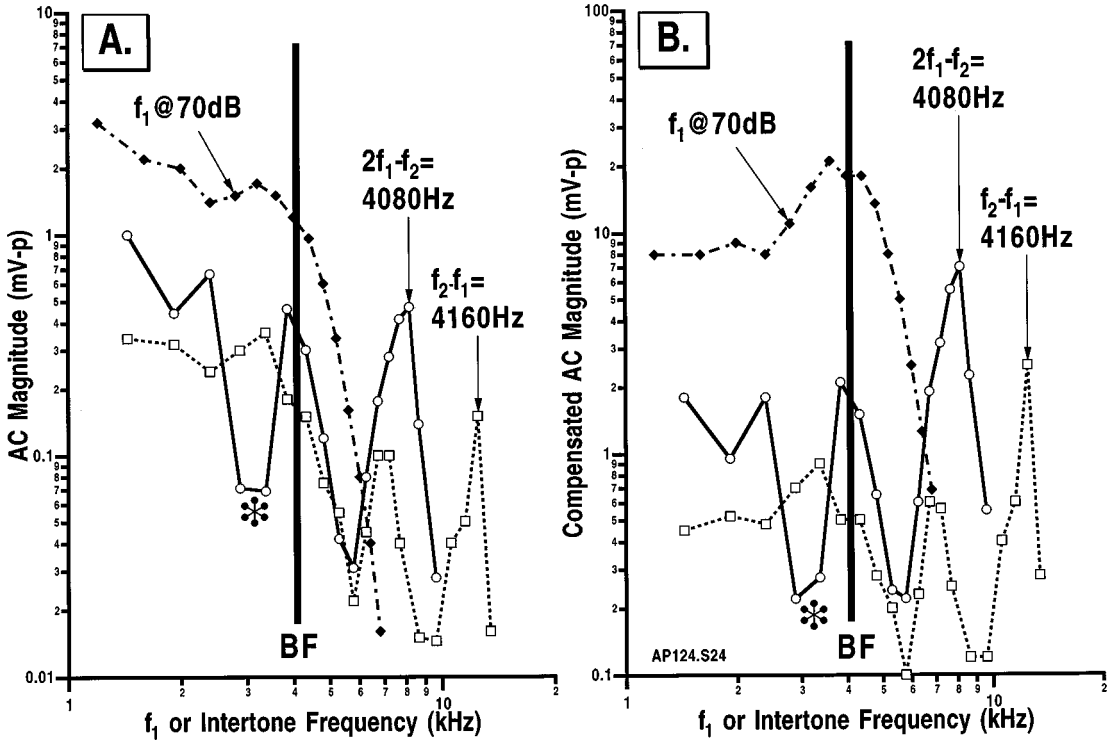


FIG. 6. These results, measured in another second-turn IHC at 70 dB, provide responses for higher-frequency primaries to assure that both the CDT and the QDT coincide with BF. A bold vertical line appears at 4000 Hz which approximates the BF in this region of the cochlea. The responses in panel (A) are peak ac voltages obtained directly from the FFT; those in panel (B), the voltages after compensating for filtering by the cell's basolateral membrane and by the recording electrode. The isolated pinwheel in each panel designates the intertone frequencies where f_2 suppresses f_1 by more than 6 dB.

B. The influence of mutual interference

Figure 6(a) provides another example from second turn where the intertone varied between 1440 and 13 440 Hz. Again the ac receptor potential for f_1 alone at 70 dB is included to indicate the response area of the IHC. The BF is estimated by the bold vertical line. As the intertone increases above BF, two peaks emerge. The first is associated with the CDT at 4080 Hz which is produced when the intertone is at 8160 Hz. Another peak appears when the QDT is near BF at 4160 Hz and the intertone is at 12 480 Hz. In Fig. 3, a traveling component was only demonstrated for the CDT because the primary pairs were not high enough to place the QDT at BF. These responses were also compensated for losses due to filtering by the cell's basolateral membrane and by the recording apparatus and plotted in panel (B) using a different ordinate. After compensation, the response to f_1 alone exhibits a peak in the BF region and the peaks for the distortion products become more prominent. Because the maximum response to both difference tones occurs around 4100 Hz, when the primaries are well above BF, it is assumed that the distortion product propagates apically from its generation site to its own BF place where it is detected by the IHC.

It is evident that the functions in Fig. 6 exhibit magnitude fluctuations. These notches are thought to reflect interactions between the two stimulating primaries as displayed in Fig. 7. In panel (A), frequency response functions at 70 dB are plotted for f_1 alone and for f_1 measured in the presence of f_2. When f_2 is at 3360 Hz, the f_1 response at 2400 Hz is suppressed to the largest extent. Companion data for f_2 are provided in panel (B). Only when f_2 is near BF at 3360 Hz does it remain unchanged in the presence of f_1. Notice that f_2 is reduced by the lower frequency primary as f_2 increases above BF. This suggests that the large decreases in distortion product magnitude shown in Fig. 6 for intertones just above BF may reflect a combination of mutual suppression in which f_1 suppresses f_2 and the fact that the primaries are moving outside of the response area of the cell. The magnitude changes are also plotted in panel (C) of Fig. 7 at the intertone to facilitate comparisons with results in Fig. 6. In fact, the pinwheel symbol associated with the large decrease in f_1 for an intertone at 2880 Hz corresponds to similar symbols included in the previous figure.

C. dc results from second turn

It is emphasized that compensated ac responses provide qualitative, not quantitative information. Consequently, it is prudent to determine the degree to which distortion products are reflected in the dc receptor potential which is not affected by filtering. The dc results in Fig. 8 are the companion measures for the ac responses shown in Fig. 6. As before, the dc frequency response function for f_1 alone at 70 dB is provided along with a reference line which approximates BF. The solid curve indicates dc responses produced when the two primaries are presented together. This two-tone function reflects a slightly broader tuning because the dc responses produced by f_1 and f_2 can sum. However, at this relatively high level, the single- and two-tone curves diverge mainly above BF where responses are not fully saturated. As the primaries increase above BF, two peaks emerge. The first

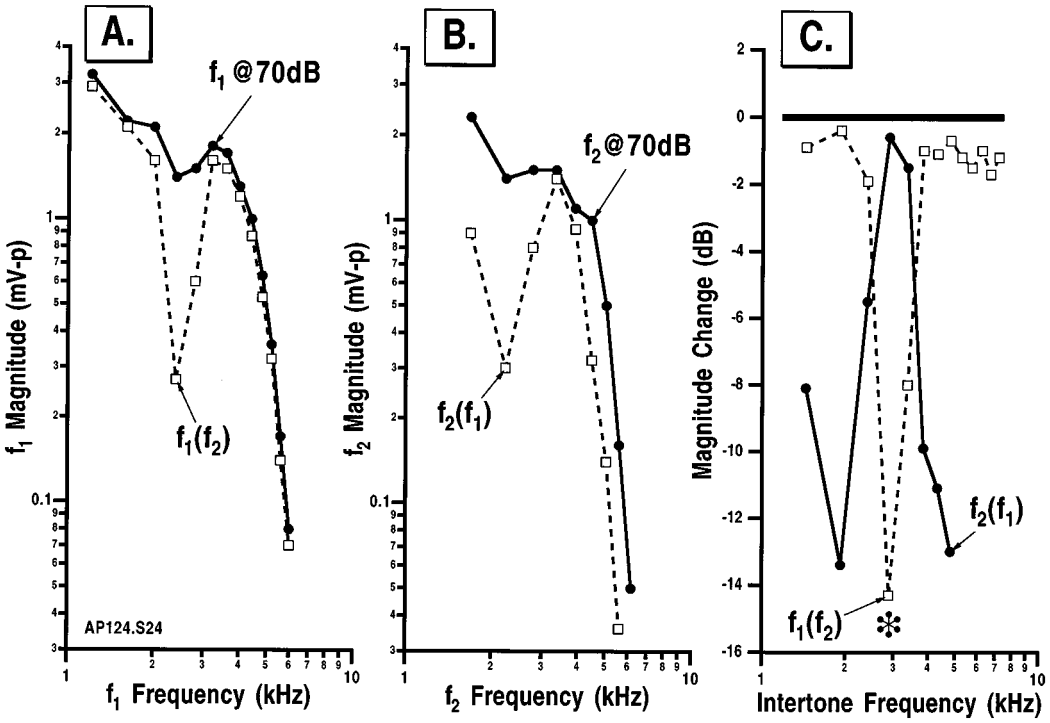


FIG. 7. Panel (A) includes frequency response functions for f_1 at 70 dB measured alone and in the presence of f_2, i.e., $f_1(f_2)$. This latter function is plotted with dashed lines. A similar presentation is provided in panel (B) for f_2 measured in the presence of f_1. Panel (C) shows the magnitude changes due to the suppression of one primary on the other. If interactions between the two inputs did not occur, then all data points would fall on the horizontal line at 0 dB. These magnitude changes are plotted at the intertone in panel (C).

occurs when the CDT is at 4080 Hz which corresponds to the ac response demonstrated in Fig. 6. In fact, the compensated ac response for this component is included and appears as a filled triangle. A second peak also emerges for higher-frequency primaries with an intertone of 12 480 Hz. This response represents the quadratic difference tone at 4160 Hz, again very near the BF of the cell. Its compensated ac response is also plotted as the open triangle.

These results indicate that distortion products with frequencies near the BF of the cell produce large dc receptor potentials even when the two stimulating primaries generate little or no response in the cell. To illustrate this point, averaged response waveforms, corresponding to these two peaks in the dc function are shown in Fig. 9. The CDT is displayed at the top for the component at 4080 Hz. Window 1 indicates the region where f_1 at 6800 Hz is presented alone; window 2, the region of overlap where the two primaries are presented together and, finally, window 3 designates the time when only f_2 at 9520 Hz is present. The large dc receptor potential in window 2 reflects the response to a CDT which is generated at a more basal location and, via its own traveling wave, is distributed to its BF place. In other words, the dc receptor potential indicated here is produced by the transducer nonlinearity of the IHC in response to ac mechanical input at $2f_1-f_2$. Because of its velocity sensitivity, the IHC does not respond to dc. At the bottom, a similar figure shows the waveform obtained when the intertone equals 12 480 Hz. Again, because the quadratic difference tone is at 4160 Hz, a dc receptor potential is recorded in the region of overlap but no response is registered to either primary alone.

D. Organ of Corti results from second turn

Responses measured in the organ of Corti fluid space for f_2-f_1 and $2f_1-f_2$ are provided in Fig. 10. These extracellular dc responses were obtained from the same animal whose intracellular IHC responses are presented in the previous figure. In both cases, the responses in window 2, the

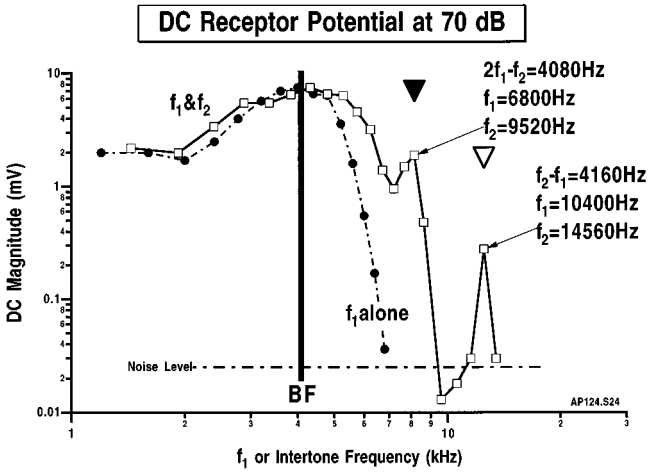


FIG. 8. This figure provides dc frequency response functions for the same cell whose ac responses are shown in Fig. 6. The function for f_1 alone is included to indicate the response area of the cell. In addition, the function obtained for the two-tone input (f_1 and f_2) is also provided and plotted at the intertone frequency. The two peaks above BF correspond to the CDT and the QDT. Both are produced by high-frequency, primary pairs. The corresponding compensated ac values are also indicated by the downward pointing triangles.

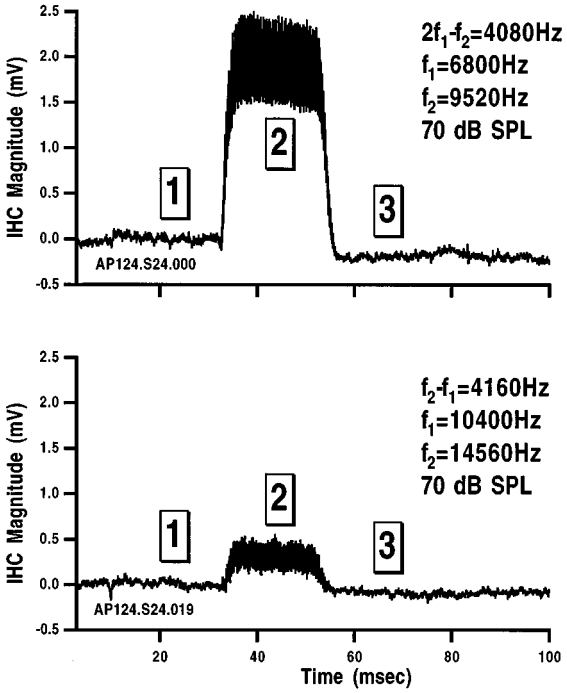


FIG. 9. The waveforms pictured here correspond to the peaks in the two-tone dc frequency response function shown in Fig. 8. That at the top is for the CDT while that at the bottom is for the QDT. The numbered boxes designate time intervals such that window 1 corresponds to the f_1 alone segment; window 2, the region of overlap, and window 3, the time when only f_2 is presented.

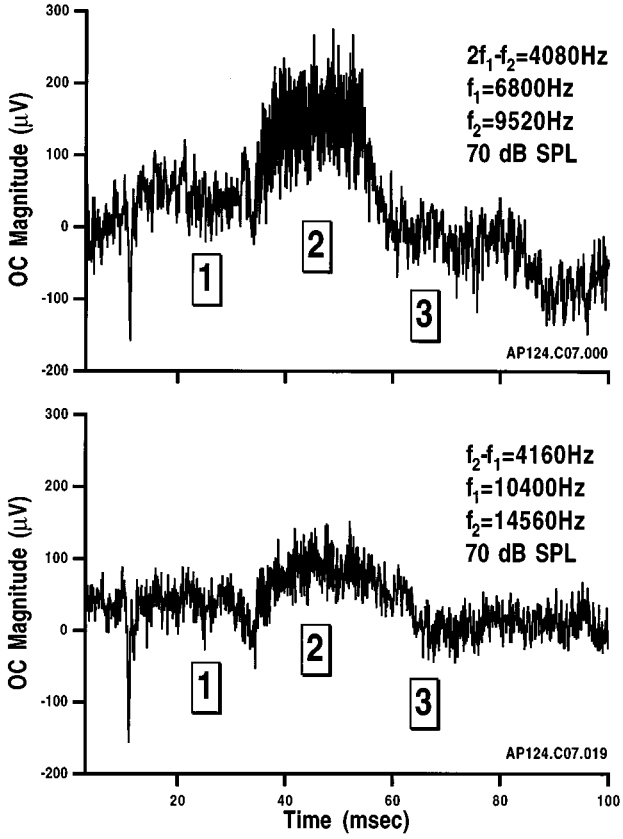


FIG. 10. Additional waveforms recorded in the organ of Corti fluid space are included here for results collected at 70 dB. These extracellular responses were obtained from the same preparation as the IHC data shown in Figs. 5–9.

region of overlap, are greater than responses to either primary alone indicating that traveling quadratic and cubic difference tones are also expressed in extracellular responses. In this example, the CDT is ~6 dB greater than the QDT. While this relationship may not hold at all levels, this magnitude disparity is smaller than that for the IHC results shown in Fig. 9. Since the data in Figs. 9 and 10 were obtained from the same animal, the magnitude differences between quadratic and cubic difference tones measured in the IHC could be overestimated. This may relate to the fact that the IHC response at f_2-f_1 was measured several minutes after that at $2f_1-f_2$. Consequently, it is possible that the organ of Corti results provide a better reflection of the relative magnitudes of individual components because the response profiles obtained here are more stable over time.

Additional data collected from the organ of Corti fluid space in second turn is provided in Fig. 11. The ac response for f_1 alone is included to indicate the broad response area of this location at 80 dB. As the intertone increases in frequency, response magnitudes for the local difference tones at f_2-f_1 and $2f_1-f_2$ decrease as expected from the shape of the f_1 alone response. However, with further increases in the intertone frequency, magnitude peaks are expressed when the distortion products are near the BF of the recording location. This occurs when the intertones are at 7200 Hz for the CDT at 3600 Hz and at 10 560 Hz for the QDT at 3520 Hz. In contrast to local components, traveling difference tones reflect narrower tuning because these distortion products, with frequencies near the BF of the recording location, are equivalent to low-level, single-tone inputs. In fact, the Q_{10} for the traveling $2f_1-f_2$ is 3.4, again consistent with the tuning of single auditory nerve fibers at the 4-kHz location in the guinea pig cochlea (Evans, 1972).

III. DISCUSSION

A. Local components generated for intertones near BF

The careful choice of stimulus parameters allows one to record locally generated distortion products when the two primaries are near the BF of the cell. As shown in Fig. 1, sum and difference tones above as well as below the primaries are prominent in ac receptor potentials recorded from individual IHCs. The magnitudes of these local components reflect the tuning of the recording location, such that distortion products with frequencies well away from BF are more severely attenuated by filtering than those closer to the BF of the cell. In addition, combination tones can be influenced by the complex interactions observed between the two stimulating primaries as in Fig. 7(C). Because the primary producing the largest response suppresses the weaker, the frequency dependence of these magnitude decreases will reflect the ways in which filter shape changes with BF and with level. In fact, at the moderately high levels used here, the largest potentials are not produced at BF because responses here are fully saturated. Frequency response functions are also broader which fosters nonlinear interactions over a wide fre-

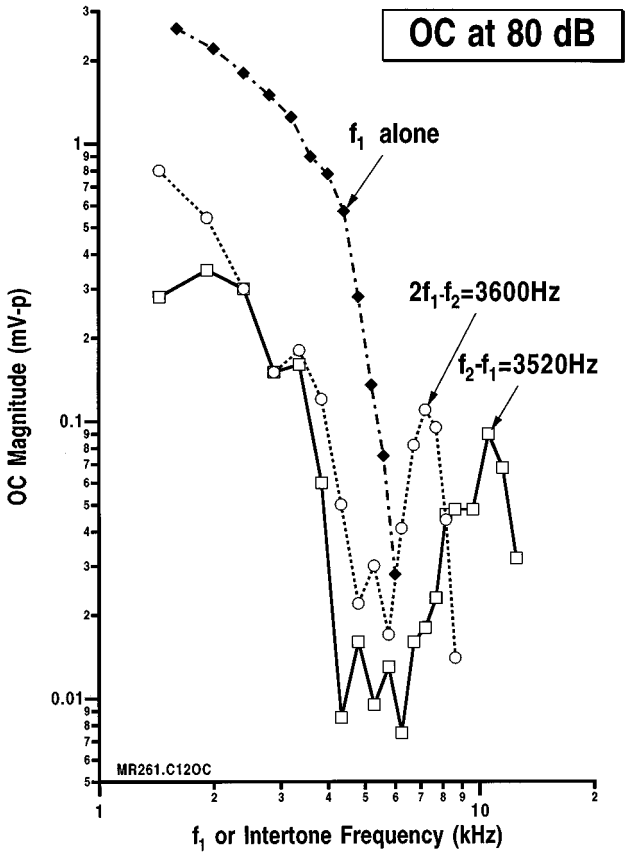


FIG. 11. The measures were made in the organ of Corti fluid space in the same general area as the second-turn IHC recordings. The voltages were obtained at 80 dB but were not compensated for reductions associated with the microelectrode.

quency range. Thus these alterations in filter shape are reflected in the frequency dependence of the suppressive interactions between the two stimulating primaries.

It should also be evident that these high-level suppression patterns contrast with the general notion that two-tone suppression is greatest around BF (Kiang and Moxon, 1974; Abbas and Sachs, 1976; Abbas, 1978; Javel *et al.*, 1978; Geisler and Sinex, 1980). This characterization, however, is largely based on single unit results obtained with a BF probe presented ~15 dB above threshold. At these low levels, tuning is sharp and the largest magnitude reductions are measured near BF. Because more linear responses are observed well above and well below BF, interactions are minimized in these regions. This general description differs from that reported here for primary pairs that increase in frequency because these measures were not obtained for low-level, BF probe tones measured in the presence of a constant-frequency suppressor. However, when suppression patterns are measured in IHCs at low-levels for BF probes, response reductions behave as indicated by the single unit results (Sellick and Russell, 1979; Cheatham and Dallos, 1989).

It should also be emphasized that since the primaries used in the present experiments were not harmonically related, magnitude notches in the two-tone functions (Fig. 6) cannot be explained by interactions between distortion products of the same frequency but of a different order (Dallos, 1973; McAnally and Calford, 1992). In fact, when frequency ratio is held constant at 1.4, fluctuations in distortion product magnitude can coincide with regions of mutual suppression where one primary suppresses the other by large amounts (Dallos and Cheatham, 1974). Suppression by the primaries on individual distortion products is also a possibility. Thus magnitude notches probably reflect a nonmonotonicity in the production of the distortion product as well as suppression upon the distortion component itself. These results imply that suppression does not follow the generation of intermodulation distortion but appears to be colocalized with the production of two-tone distortion. In other words, the same nonlinearity could be responsible for both.

Because the notches reported in the two-tone functions are influenced by mutual suppression between the two primaries, this interaction should be considered when evaluating similar response patterns reported elsewhere. These hair cell results are consistent with the idea that the presence of magnitude notches does not necessarily indicate that different sources are involved. In other words, the presence of a notch need not imply phase cancellation between two components produced by separate sources (Weiss and Leong, 1985). This conclusion, however, probably does not apply to input–output functions obtained at other frequency ratios and demonstrated in human subjects (Smoorenburg, 1972; Hall, 1975; Weber and Mellert, 1975) and in distortion product otoacoustic emissions (Weiderhold *et al.*, 1986; Brown, 1987; Zwicker and Harris, 1990; Whitehead *et al.*, 1992). These latter results probably reflect contributions from relatively large segments along the basilar membrane. Thus the notches could correspond to interactions between sources located at different positions along the cochlear partition (Furst *et al.*, 1988; Sun *et al.*, 1994a,b; Robles *et al.*, 1997; Brown *et al.*, 1996; Gaskill and Brown, 1996; Stover *et al.*, 1996). They do not necessarily imply, however, that the individual sources are different in nature.

B. Traveling components generated for intertones above BF

When primary pairs are placed well above BF, IHC responses can also be measured for distortion products with frequencies near the BF of the cell. These results are consistent with those recorded from IHCs in the base of the cochlea (Nuttall and Dolan, 1990) as well as those from the auditory nerve (Goldstein and Kiang, 1968; Kim *et al.*, 1980; Siegel *et al.*, 1982). In fact, the magnitudes of these traveling distortion products tend to exceed those produced by the primaries at the recording location. In other words, distortion products can be measured when the primaries do not produce responses in the cell when presented alone. Because these descriptions apply equally well to quadratic and cubic difference tones, they confirm the single unit results (Kim *et al.*, 1980; Siegel *et al.*, 1982) suggesting that both components have their own traveling waves. In other words, that both f_2-f_1 and $2f_1-f_2$ are generated at more basal locations, in the region of overlap between the two stimulating primaries, and are distributed to their own BF place.

Although similar behavior has been observed in basilar membrane mechanics in the base of the cochlea for the CDT, a QDT has not yet been measured (Robles *et al.*, 1991, 1997;

Nuttall and Dolan, 1993). These mechanical experiments, however, utilized primary pairs which produced a CDT near BF but a QDT well below BF. Thus the absence of a component at f_2-f_1 may reflect the fact that frequency combinations favorable to the QDT have not been used. This is important because only when distortion products are within the passband of the location under study is a response observed. This holds for both local and traveling components.

This situation is similar in some ways to that in psychophysics where QDTs and CDTs were thought at one time to exhibit different characteristics and to be produced by different sources (Zwicker, 1955, 1979; Goldstein, 1967). However, when both components were examined for primaries at similar input levels, the two difference tones behaved in a similar fashion (Hall, 1972; Humes, 1985; Zwicker and Martner, 1990). Consequently, one should not rule out the possibility that they are produced by a single source which is consistent with the IHC results reported here.

Although *in vivo* measurements are not available from OHCs, data from the organ of Corti fluid space are relevant. This is because the cochlear microphonic is thought to reflect ac receptor potentials produced by nearby OHCs (Dallos and Cheatham, 1976). In addition, the CM is thought to mirror the displacement pattern of the basilar membrane at least for inputs below BF (Dallos *et al.*, 1974). Consequently, the QDT exhibited in the organ of Corti ac response suggests that further mechanical experiments are warranted. Use of higher-frequency primaries and larger frequency ratios should result in a QDT with a frequency nearer to the site of observation along the basilar membrane. If IHC and organ of Corti results do generalize to the base of the cochlea, then a traveling QDT may also be expressed in cochlear mechanics.

It could be argued that a QDT will not be measured in high-frequency regions because the set point of OHCs, in contrast to IHCs and OHCs in the apex of the cochlea (Dallos and Cheatham, 1992), is located at the place of maximum gain on the transducer operating curve (Cody and Russell, 1987; Frank and Kössl, 1996). If true, then cubic rather than quadratic components should dominate. However, for the organ of Corti data reported in Figs. 10 and 11, traveling QDTs were produced by primaries with intertones of 12 480 and 10 560 Hz. Thus, these data suggest that quadratic components *are* produced in the 10–12 kHz region of the guinea pig cochlea implying asymmetrical transfer functions.

IV. CONCLUSION

Results in this report indicate that intermodulation distortion components are prominent in the generation region when the intertone is placed near the BF of cochlear hair cells. It is also shown that difference tones at f_2-f_1 and $2f_1-f_2$ produce IHC receptor potentials even when responses are not observed for either of the two stimulating primaries. This finding is consistent with single unit results (Kim *et al.*, 1980; Siegel *et al.*, 1982) and suggests that both quadratic and cubic difference tones have their own traveling waves. It was also observed that mutual suppression between the primaries, and possibly between the primaries and individual distortion products, influences the magnitudes of combination tones at the site of generation. Thus, these interactions should be considered when evaluating whether quadratic and cubic components are produced by separate, nonlinear sources.

ACKNOWLEDGMENTS

This work was supported by research Grant No. 5 R01 DC00089 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health.

Abbas, P. (**1978**). ''Effects of stimulus frequency on two-tone suppression,'' J. Acoust. Soc. Am. **63,** 1878–1886.

Abbas, P., and Sachs, M. (**1976**). ''Two-tone suppression in auditory nerve fibers: Extension of a stimulus-response relationship,'' J. Acoust. Soc. Am. **59,** 112–122.

Baden-Kristensen, K., and Weiss, T. F. (**1983**). ''Receptor potentials of lizard hair cells with free-standing stereocilia: Responses to acoustic clicks,'' J. Physiol. (London) **335,** 699–721.

Brown, A. M. (**1987**). ''Acoustic distortion from rodent ears: A comparison of responses from rats, guinea pigs and gerbils,'' Hearing Res. **31,** 25–38.

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.

Cheatham, M. A., and Dallos, P. (**1989**). ''Two-tone suppression in inner hair cell responses,'' Hearing Res. **40,** 187–196.

Cheatham, M. A., and Dallos, P. (**1992**). ''Physiological correlates of off-frequency listening,'' Hearing Res. **59,** 39–45.

Cheatham, M. A., and Dallos, P. (**1993**). ''Longitudinal comparisons of IHC ac and dc receptor potentials recorded from the guinea pig cochlea,'' Hearing Res. **68,** 107–114.

Cody, A. R., and Russell, I. J. (**1987**). ''The properties of hair cells in the basal turn of the guinea-pig cochlea to tones,'' J. Physiol. **383,** 551–569.

Dallos, P. (**1969**). ''Combination tone $2f_1-f_h$ in microphonic potentials,'' J. Acoust. Soc. Am. **46,** 1437–1444.

Dallos, P. (**1973**). *The Auditory Periphery* (Academic, New York).

Dallos, P. (**1977**). ''Comment to W. S. Rhode,'' in *Psychophysics and Physiology of Hearing*, edited by E. F. Evans and J. P. Wilson (Academic, New York), p. 39.

Dallos, P. (**1984**). ''Some electrical circuit properties of the organ of Corti. II. Analysis including reactive elements,'' Hearing Res. **14,** 281–291.

Dallos, P. (**1985**). ''Response characteristics of mammalian cochlear hair cells,'' J. Neurosci. **5,** 1591–1608.

Dallos, P., and Cheatham, M. A. (**1974**). ''Cochlear microphonic correlates of cubic difference tones,'' in *Facts and Models in Hearing*, edited by E. Zwicker and E. Terhardt (Springer-Verlag, New York), pp. 312–322.

Dallos, P., and Cheatham, M. A. (**1976**). ''Production of cochlear potentials by inner and outer hair cells,'' J. Acoust. Soc. Am. **60,** 510–512.

Dallos, P., and Cheatham, M. A. (**1992**). ''Cochlear hair cell function reflected in intracellular recording in vivo,'' in *Sensory Transduction*, edited by D. P. Corey and S. D. Roper (Society of General Physiologists Symposium, Rockefeller U. P., New York), pp. 371–393.

Dallos, P., Cheatham, M. A., and Ferraro, J. (**1974**). ''Cochlear mechanics, nonlinearities and cochlear potentials,'' J. Acoust. Soc. Am. **55,** 597–605.

Dallos, P., Santos-Sacchi, J., and Flock, Å. (**1982**). ''Intracellular recordings from cochlear outer hair cells,'' Science **218,** 582–584.

Dallos, P., Harris, D. M., Relkin, E., and Cheatham, M. A. (**1980**). ''Two-tone suppression and intermodulation distortion in the cochlea: Effects of outer hair cell lesions,'' in *Psychological, Physiological and Behavioral Studies in Hearing*, edited by G. van den Brink and F. A. Bilsen (Delft U. P., Noordwijkerhout, The Netherlands), pp. 242–252.

Evans, E. F. (**1972**). ''The frequency response and other properties of single fibres in the guinea-pig cochlear nerve,'' J. Physiol. (London) **226,** 263–287.

Frank, G., and Kössl, M. (**1996**). ''The acoustic two-tone distortions $2f_1-f_2$ and f_2-f_1 and their possible relation to changes in the operating point of the cochlear amplifier,'' Hearing Res. **98,** 104–115.

Furst, M., Rabinowitz, W. M., and Zurek, P. M. (**1988**). ''Ear canal acoustic distortion at $2f_1-f_2$ from human ears: Relation to other emissions and perceived combination tones,'' J. Acoust. Soc. Am. **84,** 215–221.

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.

Geisler, C. D., and Sinex, D. G. (**1980**). ''Responses of primary auditory fibers to combined noise and tonal stimuli,'' Hearing Res. **3,** 317–334.

Gibian, G. L., and Kim, D. O. (**1979**). ''Cochlear microphonic evidence for mechanical propagation of distortion products (f_2-f_1) and $(2f_1-f_2)$,'' J. Acoust. Soc. Am. Suppl. 1 **65,** S84.

Gibian, G. L., and Kim, D. O. (**1982**). ''Cochlear microphonic evidence for mechanical propagation of distortion products (f_2-f_1) and $(2f_1-f_2)$,'' Hearing Res. **6,** 35–59.

Goldstein, J. L. (**1967**). ''Auditory nonlinearity,'' J. Acoust. Soc. Am. **41,** 676–689.

Goldstein, J. L., and Kiang, N. Y.-S. (**1968**). ''Neural correlates of the aural combination tone $2f_1-f_2$,'' Proc. IEEE **56,** 981–992.

Guinan, J., and Peake, W. T. (**1967**). ''Middle ear characteristics of anesthetized cats,'' J. Acoust. Soc. Am. **41,** 1237–1261.

Hall, J. L. (**1972**). ''Auditory distortion products (f_2-f_1) and ($2f_1-f_2$),'' J. Acoust. Soc. Am. **51,** 1863–1871.

Hall, J. L. (**1975**). ''Nonmonotonic behavior of distortion product $2f_1-f_2$: Psychophysical observations,'' J. Acoust. Soc. Am. **58,** 1046–1050.

Helmholtz, H. L. F. (**1954**). *On the Sensations of Tone* (Dover, New York).

Humes, L. E. (**1985**). ''Cancellation level and phase of the (f_2-f_1) distortion product,'' J. Acoust. Soc. Am. **78,** 1245–1251.

Javel, E., Geisler, C. D., and Ravindran, A. (**1978**). ''Two-tone suppression in auditory nerve of the cat: Rate-intensity and temporal analysis,'' J. Acoust. Soc. Am. **63,** 1093–1104.

Kiang, N. Y.-S., and Moxon, E. C. (**1974**). ''Tails of tuning curves of auditory nerve fibers,'' J. Acoust. Soc. Am. **55,** 620–630.

Kim, D. O., Molnar, C. C., and Matthews, J. W. (**1980**). ''Cochlear mechanics: nonlinear behavior in two-tone responses as reflected in cochlear-nerve fiber responses and in ear-canal sound pressure,'' J. Acoust. Soc. Am. **67,** 1704–1721.

Lewis, D., and Reger, S. N. (**1933**). ''An experimental study of the role of the tympanic membrane and the ossicles in the hearing of certain subjective tones,'' J. Acoust. Soc. Am. **5,** 153–158.

McAnally, K. I., and Calford, M. B. (**1992**). ''Coherence of frequency modulation is encoded by cochlear-generated distortion,'' Hearing Res. **58,** 213–220.

Nuttall, A. L., and Dolan, D. F. (**1990**). ''Inner hair cell responses to the $2f_1-f_2$ intermodulation distortion product,'' J. Acoust. Soc. Am. **87,** 782–790.

Nuttall, A. L., and Dolan, D. F. (**1993**). ''Intermodulation distortion (f_2-f_1) in inner hair cell and basilar membrane responses,'' J. Acoust. Soc. Am. **93,** 2061–2068.

Nuttall, A. L., Dolan, D. F., and Avinash, G. (**1990**). ''Measurements of basilar membrane tuning with laser Doppler velocimetry,'' in *The Mechanics and Biophysics of Hearing*, edited by P. Dallos, C. D. Geisler, J. W. Matthews, M. A. Ruggero, and C. R. Steele (Springer-Verlag, New York), pp. 288–295.

Ohm, G. S. (**1843**). ''Ueber die definition des tones nebst daran geknüpfter theorie der sirene and ähnlicher tonbildenden vorrichtungen,'' Ann. Phys. Chem. **59**, 513–565.

Patuzzi, R., Yates, G. K., and Johnstone, B. M. (**1989**). ''Outer hair cell receptor current and sensorineural hearing loss,'' Hearing Res. **42,** 47–72.

Robles, L., Ruggero, M. A., and Rich, N. C. (**1991**). ''Two-tone distortion in the basilar membrane of the cochlea,'' Nature (London) **349,** 413–414.

Robles, L., Ruggero, M. A., and Rich, N. C. (**1997**). ''Two-tone distortion at the basilar membrane of the chinchilla cochlea,'' J. Neurophysiol. **77,** 2385–2399.

Sellick, P. M., and Russell, I. J. (**1979**). ''Two-tone suppression in cochlear hair cells,'' Hearing Res. **1,** 227–236.

Siegel, J. H., Kim, D. O., and Molnar, C. E. (**1982**). ''Effects of altering organ of Corti on cochlear distortion products f_2-f_1 and $2f_1-f_2$,'' J. Neurophysiol. **47,** 303–328.

Smoorenburg, G. F. (**1972**). ''Combination tones and their origin,'' J. Acoust. Soc. Am. **52,** 615–632.

Stover, L. J., Neely, S. T., and Gorga, M. P. (**1996**). ''Latency and multiple sources of distortion product otoacoustic emission,'' J. Acoust. Soc. Am. **99,** 1016–1024.

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.

Weber, R., and Mellert, V. (**1975**). ''On the nonmonotonic behavior of cubic distortion products in the human ear,'' J. Acoust. Soc. Am. **57,** 207–214.

Weiss, T. F., and Leong, R. (**1985**). ''A model for signal transmission in an ear having hair cells with free-standing stereocilia. IV. Mechanoelectric transduction stage,'' Hearing Res. **20,** 175–195.

Wever, E. G., Bray, C. W., and Lawrence, M. (**1940**). ''A quantitative study of combination tones,'' J. Exp. Psychol. **27,** 469–496.

Whitehead, M. L., Lonsbury-Martin, B. L., and Martin, G. K. (**1992**). ''Evidence for two discrete sources of $2f_1-f_2$ distortion-product otoacoustic emission in rabbit: I. Differential dependence on stimulus parameters,'' J. Acoust. Soc. Am. **91,** 1587–1607.

Wiederhold, M. L., Mahoney, J. W., and Kellogg, D. L. (**1986**). ''Acoustic overstimulation reduces $2f_1-f_2$ cochlear emissions at all levels in the cat,'' in *Peripheral Auditory Mechanisms*, edited by J. B. Allen, J. L. Hall, A. E. Hubbard, S. T. Neely, and A. Tubis (Springer-Verlag, New York), pp. 322–329.

Zwicker, E. (**1955**). ''Der ungewöhnliche amplitudengang der nichtlinearen verzerrungen des ohres,'' Acustica **5,** 67–74.

Zwicker, E. (**1979**). ''Different behavior of quadratic and cubic difference tones,'' Hearing Res. **1,** 283–292.

Zwicker, E., and Harris, F. P. (**1990**). ''Psychophysical and ear canal cancellation of $(2f_1-f_2)$-distortion products,'' J. Acoust. Soc. Am. **87,** 2583–2591.

Zwicker, E., and Martner, O. (**1990**). ''On the dependence of (f_2-f_1) difference tones on subject and on additional masker,'' J. Acoust. Soc. Am. **88,** 1351–1358.

Otoacoustic emissions, hair cells, and myosin motors

Geoffrey A. Manley[a] and Lothar Gallo
Institut für Zoologie, Technische Universität München, Lichtenbergstr. 4, 85747 Garching, Germany

(Received 17 December 1996; revised 3 March 1997; accepted 4 March 1997)

The stereovillar bundles of hair cells show active movements that may be generated by the putative myosin–actin interactions underlying hair-cell adaptation. Such movement is a possible candidate for the generation of spontaneous otoacoustic emissions (SOAE) in the ear canal of nonmammals. In the basilar papilla of certain lizard families, most hair cells are not coupled by a tectorial membrane, making it easier to assign the energy in emission peaks to defined groups of hair cells. We have studied 62 SOAE in the Bahamian Anole *Anolis sagrei*, which has about 140 hair cells with ''free-standing'' bundles in the high-frequency area of its papilla. Individual SOAE peaks were traced to between 3 and 38 hair cells, and the mean power output per hair cell was calculated to be 141 aW. The number of bundle myosins putatively involved in the generation of each SOAE was estimated and the force generated by each myosin at 1 kHz calculated to be approximately 0.1 pN. The data support the idea that hair cells generate emissions and suggest that myosin produces sufficient power to be the emission motor. © *1997 Acoustical Society of America.*
[S0001-4966(97)03007-5]

PACS numbers: 43.64.Jb, 43.64.Kc [RDF]

INTRODUCTION

Spontaneous otoacoustic emissions (SOAE) are sounds in the external ear canal that result from mechanical energy output of the receptor cells, or hair cells, of the hearing organ (review in Köppl, 1995). This close association is demonstrated, e.g., by the great similarity between the frequency selectivity of suppression phenomena for SOAE on the one hand and single neural tuning curves on the other (Köppl, 1995; Köppl and Manley, 1994; Zizz and Glattke, 1988). In mammals, it is generally assumed that outer hair cells (OHC) are specialized to produce rapid length changes that couple back in some as yet unknown way into the motion of the epithelium; they are thus viewed as the force generators underlying both the cochlear amplifier and the generation of SOAE (e.g., Dallos and Evans, 1995). SOAE have been described in all groups of tetrapods, however (for review see Köppl, 1995), and the mechanism proposed to underly their generation in mammals involves special structural and molecular adaptations of the outer hair cells that have not been demonstrated in nonmammals. Motility of the hair-cell body, as observed in mammalian OHC, is unlikely to apply to nonmammalian hearing organs (Klinke and Smolders, 1993; Köppl, 1995; Manley, 1995).

Another mechanism presently under consideration as a source of mechanical energy is the hair-cell bundle itself, that has been shown to undergo active movements (Crawford and Fettiplace, 1985; Denk and Webb, 1992; Howard and Hudspeth, 1988), including rapid, biphasic twitches during deflection and even spontaneous twitching (Benser *et al.*, 1996). Movement of the bundle is also associated with the gating springs that are hypothesized to be essential for the function of the hair-cell transduction channels, including their adaptation (Assad and Corey, 1992; Howard *et al.*, 1988; Hudspeth and Gillespie, 1994). The gating-spring hypothesis (Hudspeth and Gillespie, 1994) proposes that every hair-cell stereovillus has at least one transduction channel connected to an elastic element that plays an essential role in sensory transduction. Transduction channels are probably activated by the stretch of tip links between the stereovilli (Pickles *et al.*, 1984) and the tip links are probably the elastic elements of the gating springs. An insertional plaque joins the upper end of each tip link to the actin core of the next-taller stereovillus and contains an estimated number of 50 Type 1 myosin molecules (Gillespie *et al.*, 1993; Hudspeth and Gillespie, 1994). In unstimulated hair cells, these myosins climb up the actin and maintain the steady tension on the tip link that keeps an average of 15% of the transduction channels in an open state (Hudspeth and Gillespie, 1994). These myosins are proposed to underly the mechanism of adaptation of hair-cell bundles, restoring the resting tension when the bundle is subjected to a steady displacement. This is achieved by a passive sliding of the plaque down the actin to compensate for a tip-link tension that is too high and by the plaque actively climbing up the actin to restore tension when it is too low (Assad and Corey, 1992). Removal of Ca^{2+} ions leads to tip-link breakage and thus release of tension between the stereovilli. After breakage, the bundle relaxes and moves 120 nm toward the tallest stereovilli and the cell ceases to transduce (Assad *et al.*, 1991).

The hypothesis we propose assumes the direct involvement of the adaptational myosin motors in active and fast hair-cell bundle movements, producing SOAE in nonmammals (Manley *et al.*, 1995; Benser *et al.*, 1996). Myosin has also previously been implicated as a candidate molecule to participate in mechanical amplification (Hudspeth and Gillespie, 1994). The rate of movement of the tip-link insertional plaques during adaptation has been measured at 2 μm/s, which is comparable to the speed of myosin motion in muscle (Hudspeth and Gillespie, 1994). We (Manley *et al.*,

[a]Corresponding author. Electronic mail: GAM@CIP1.zoo.chemie.tu-muenchen.de

1995) and others (Hudspeth and Gillespie, 1994) have previously suggested that the myosins in the adaptational plaques may be stretch activated, as are the myosins in the indirect flight muscles of insects (Tregear, 1975).

In insects, there are two basic types of flight-muscle control systems. In large insects, flight muscles receive a nerve impulse for each cycle of the wing beat; this mechanism supports relatively low wing-beat frequencies. In addition to considerations of size effects, this so-called myoneural synchrony is limited by the muscle cells' ability to rapidly pump the calcium ions necessary for activating and stopping the contraction cycle. Calcium pumping rates would limit the wing-beat frequency to about 100 Hz (Sparrow, 1995). This problem is eliminated in the second group of flying insects that have indirect (or asynchronous) control of their flight muscles. In these animals, the neural input to the muscles is not synchronous with contractions, but much slower. The function of the neural input is to set up an active state in which during flight, enough calcium is always present in the cytoplasm to support myosin–actin interaction. Following the setting-up of an active state in the flight muscles, the insect actually begins flight when a special starter muscle contracts and stretches one of the blocks of flight muscle. In the active state, stretch leads to automatic contraction (of only 1%–2% of muscle length), and this contraction pulls on the thorax wall in such a way that the second muscle block is stretched, and this then reflexively contracts (Tregear, 1975; Sparrow, 1995). The reciprocal stretching continues as long as the muscles are in the active state (i.e., the muscle fiber is depolarized by neural input). The frequency of oscillation of these muscles is determined by the mechanics of the wing-thorax system and is not limited by the speed of the muscles (Tregear, 1975).

Our hypothesis is that a similar stretch activation of the myosins in the adaptational plaques of hair-cell stereovilli could sustain active oscillations in micromechanically tuned hair-cell bundles (Manley *et al.*, 1995). The equivalent of the ''active state'' as in insect flight muscle might be induced by the steady ion flow through open transduction channels. Even simple rocking back-and-forth of the head regions of myosin molecules (with a stroke of 11 nm) could, in sacculus hair cells, move the bundle by up to 80 nm, which is many times greater than stimulus-threshold displacement values (Benser *et al.*, 1996; Hudspeth and Gillespie, 1994). It has been suggested that in asynchronous insect flight muscles, that only contract by about 1%–2% of their length, the myosin cross-bridges remain attached during the movement cycle (Alexander, 1995). Benser *et al.* (1996) believe it is plausible that in hair cells, the insertional plaque myosins respond to a stretch stimulus by tilting the head region of the molecule. The goal of the present research was to begin to test this myosin hypothesis by asking whether the properties of SOAE are compatible with it. In this paper, we assess whether the power emitted during the production of SOAE is compatible with an origin in myosin generators.

SOAE are generally found as clear peaks in frequency spectra of ear-canal sound pressure, peaks that differ in center frequency, frequency bandwidth, and peak spectral level (Köppl, 1995). One of the problems associated with studying the force generation by hair cells underlying SOAE is to know just how many hair cells are generating individual emission peaks. Due to coupling phenomena, especially through the tectorial membrane, it was to date impossible to know the number of hair cells involved, and this has precluded estimates of the power generated per hair cell. We have pursued this question in a species of lizard in which a tectorial membrane is absent in the hair-cell area generating SOAE and in which the hair-cell bundles are thus only fluid coupled (Manley, 1990; Wever, 1978; Miller, 1992).

I. MATERIALS AND METHODS

A. SOAE measurements

SOAE were measured near the eardrum of the iguanid lizard *Anolis sagrei* in a closed acoustic system using a probe consisting of a $\frac{1}{2}$-in. condensor microphone (B&K type 4166) in a custom-made short, low-volume cone that connected to the animal's outer ear canal. The probe's impedance was optimized to permit reliable measurements even from higher-frequency peaks. The microphone was connected to a low-noise preamplifier (B&K Type 2660, +20 dB) and fed into a dynamic spectrum analyzer (Stanford SR 760 FFT). Averaged rms spectra (Hanning window) were stored on disk and analyzed off-line. In general, 200 spectra were averaged.

Two steps were necessary in order to estimate the power being generated by a given number of hair cells. In the first step, the power in each SOAE peak had to be estimated. In the second step, the anatomy of the hearing organ was investigated and the frequency map deduced, which could then be used to estimate the number of hair cells producing individual SOAE peaks.

B. Estimation of the power in each SOAE peak

The power in each spectral peak was calculated according to the following rationale: Assuming that in the ear canal the sound pressure and velocity are in phase, the sound power P can be calculated from the level L in the following steps (Zwicker and Fastl, 1990). Intensity I is related to level L by the equation

$$L = 10 \log \frac{I}{I_0} \text{ dB},$$

where I_0 is the internationally standardized reference intensity of 1 pW m^{-2}. Rewriting this equation for the intensity I corresponding to a level L, we obtain

$$I = 10^{L/10} \cdot I_0 .$$

Sound power P is the product of intensity I and the surface area S. Hence for the relationship between level L and the sound power P, we obtain

$$P = I \cdot S = 10^{L/10} \cdot I_0 \cdot S.$$

For example, for a level of $L = 0$ dB SPL, and the surface area of the $\frac{1}{2}$-in. measuring microphone $S = 1.27$ cm^2, we obtain

$$P = 10^{0/10} \text{ pW m}^{-2} \cdot 1.27 \text{ cm}^2 = 1.27 \cdot 10^{-16} \text{ W} = 0.127 \text{ fW}.$$

If sound pressure and velocity are not always in phase, the values we measure would be too small. Within a given SOAE peak, the power in all spectral lines (linewidth 15.6 Hz) was summed. To calculate how much power is present in the fluids within the hearing organ, we assumed that the middle ear is essentially an ideal transformer (Magnan *et al.*, 1996) that shows only frictional and resistive losses of 12 dB (e.g., Manley and Johnstone, 1974). Thus we assumed that the amount of power measurable in the ear canal is reduced 12 dB compared to the available power in the inner-ear fluids.

C. Anatomy of the basilar papilla

Four basilar papillae were fixed *in situ* by submersion of the whole head in a cooled solution of 2.5% glutaraldehyde and 15% picric acid overnight. The cochlear duct was carefully dissected free and the vestibular membrane opened in a 70% ethanol solution. The papillae were then taken through an alcohol series up to 100%, transferred to acetone, and critical-point dried in CO_2. Following gold sputtering, the papillae were photographed at various magnifications in a scanning electron microscope (Jeol JSM-IC 25S). From the photographs, measurements were made of the papillar length and of the heights of bundles in those cases where the tallest stereovilli were visible at appropriate angles. In two cases, the hair-cell bundles were ''shaved'' off the surface of the papilla using an eyelash and the papillae were recoated with gold. In these cases, the number, position, and surface area of the hair cells was measured and the number of stereovillar ''stumps'' on each hair cell was counted.

D. Derivation of the papillar frequency map

It was necessary to estimate the frequency map of the papilla in order to attribute the frequencies in particular emission peaks to a certain group of hair cells. We assumed that, as in other lizards (Manley, 1990; Köppl, 1995), frequencies above 1 kHz are mapped logarithmically along the bidirectionally oriented hair-cell area. The highest frequency on the papilla was assumed to be the next whole kHz to the SOAE with the highest center frequency, which was 7.7 kHz. Thus we assumed that frequencies between 1 and 8 kHz are coded on a logarithmic scale along the 250 μm length of the high-frequency region of the *Anolis* papilla [Fig. 1(b)].

E. Estimation of the number of hair cells generating each SOAE peak

The frequency map assumed in Sec. I D was projected graphically onto scanning EM photographs of the papilla and the number of hair cells within the spectral limits of each SOAE peak were counted. We assumed all hair cells within the limits of each spectral peak were (a) entrained to the same frequency and (b) contributed equally to generating the peak. We also assumed that even though the bundle height changes along the papilla, the bundle motion couples equally well into the fluid. Our estimate is thus an upper bound for the number of hair cells involved.

All the work described here was carried out in accordance with German animal-protection law.

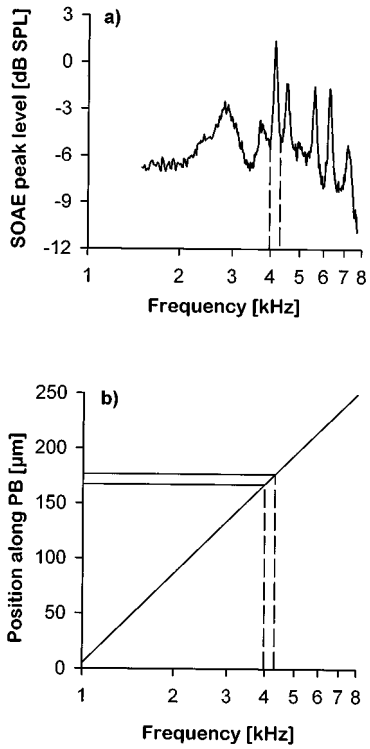


FIG. 1. (a) A typical averaged acoustic spectrum (200 averages) from the ear canal of *Anolis sagrei*. The assumed frequency limits of one of the emission peaks are shown as vertical dashed lines to the frequency axis. (b) The frequency limits read off as in (a) were projected onto an estimated frequency map for the high-frequency area of the papilla (i.e., the area of ''free-standing'' hair-cell bundles) to obtain the positional limits of the hair-cell area responsible for this individual peak on the length axis. From scanning electron-microscope images of the papilla, we counted the number of hair cells falling between—in this case—the positions 165 and 175 μm from the end of that part of the hair-cell area made up of hair cells with free-standing bundles.

II. RESULTS

A. Anatomy of the papilla

The papilla of *Anolis sagrei* contains about 140 hair cells in the bidirectionally oriented hair-cell area, which is 250 μm long. Since the highest SOAE center frequency lay at 7.7 kHz, we assumed that this area of the papilla analyzed frequencies between 1.0 and 8.0 kHz, a range of three octaves. The hair-cell density as estimated for each octave changed from 2.5 hair cells/100 μm^2 for the octave between 1 and 2 kHz, 2.9 hair cells/100 μm^2 from 2 to 4 kHz, and 4.1 hair cells/100 μm^2 between 4 and 8 kHz. All stereovillar bundles in the high-frequency hair-cell area were free-standing and the number of stereovilli per hair-cell bundle ranged from 51 in the 1-kHz region to 62 in the 8-kHz region. The stereovilli of the shortest row, which have no top ends of tip links and therefore no myosin motors, were not taken into account in calculating the number of myosin molecules; approximately 11 stereovilli per hair cell are in the shortest row. Along the relevant part of the papilla, the height of the tallest stereovilli in each bundle fell from about 20 μm to 3 μm.

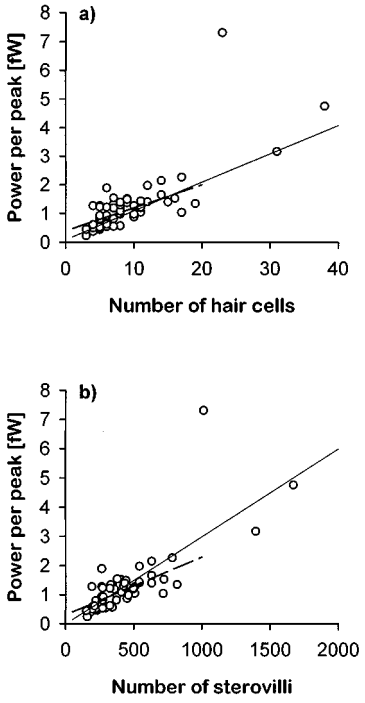


FIG. 2. (a) The power per emission peak for all SOAE as a function of the number of hair cells involved; the continuous line is a linear fit ($n=62$, $r=0.764$, $p<0.001$). The power per peak approximately doubles with a doubling of the number of hair cells. The few (3) very broad emissions have only a small influence on the slope of the fit. Even without these broad emissions, the correlation is statistically significant (dashed line, $n=59$, $r=0.671$, $p<0.001$). (b) Since the number of stereovilli per hair cell along the high-frequency area only changes by about 20%, the power output also rises with the number of stereovilli that are potentially involved. The continuous and dashed lines represent linear fits with and without the three broad emissions, as in (a); ($r=0.77$ and 0.706, respectively; $p<0.001$ in both cases).

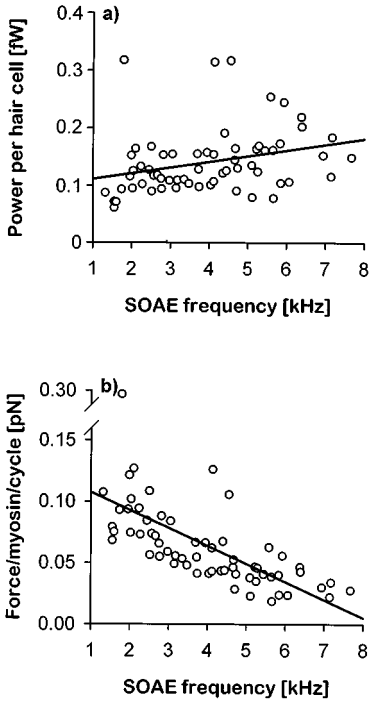


FIG. 3. (a) The calculated power output per hair cell as a function of the center frequency of the emission peaks. The line is a linear regression fit ($n=62$, $r=0.297$, $p<0.02$). (b) The estimated force produced per myosin molecule per cycle as a function of the center frequency of the SOAE. The line is a linear regression fit ($n=62$, $r=-0.602$, $p<0.001$).

B. Estimation of the number of hair cells generating each SOAE peak

We measured 62 SOAE spectral peaks in 9 ears of 6 animals (all at 30 °C). An example of an averaged spectrum is given in Fig. 1(a). The center frequencies of the SOAE peaks ranged from 1.3 to 7.7 kHz. By measuring the frequency bandwidth of each peak at its base and projecting these frequency limits on to the papilla's putative frequency map [Fig. 1(b)], the 62 emission peaks were traced to between 3 and 38 hair cells. The calculated power per emission peak was found to be clearly correlated with the estimated number of hair cells involved [Fig. 2(a)]. The power output also rose as a function of the potential number of stereovilli involved [Fig. 2(b)]. This is a consequence of the correlation in Fig. 2(a), since there is only a small change (20%) in the number of stereovilli per hair cell along the high-frequency area. It might be assumed that the correlation between peak power and number of hair cells is simply determined by the fact that there is a logarithmic frequency distribution on the papilla. The power per peak is, however, not only well correlated with the peak width. It correlated equally well with peak level.

C. Estimation of the power generated by individual myosin molecules

We estimated the power output for individual myosin molecules to compare this estimate to previously published figures for myosin power. From the total number of contributing stereovilli in each case, we estimated how many myosin motors could be responsible for individual SOAE spectral peaks. This calculation assumes: (1) that the myosin plaque for each tip link contains 50 myosins (Gillespie *et al.*, 1993; Hudspeth and Gillespie, 1994); (2) that only 10% of the myosins in any given tip-link plaque are synchronously active, as suggested by Hudspeth and Gillespie (1994; this is equivalent to assuming that each myosin molecule does not add a force contribution to every cycle of the bundle oscillation, but on average only to every tenth cycle); and (3) that as in muscle cells, the myosin molecule's cross-bridge region moves 11 nm per active force-generating cycle (Huxley, 1990; Finer *et al.*, 1994). It was not necessary to assume that the myosin head couples and uncouples during this movement. Knowing the frequencies of oscillation involved in each SOAE, i.e., using the center frequency of the SOAE peak, the force exerted per myosin molecule per cycle was calculated for each emission.

The calculated power output per hair cell showed a small dependence on frequency [Fig. 3(a)] that was, however, not significant ($r=0.297$, $p<0.02$). The force generated per cycle by each myosin fell systematically with rising frequency [Fig. 3(b)]. In the case of the present data, the mean power output per hair cell was calculated to be 141 aW, and the average force generated per myosin and per

cycle was 0.064 pN. Our calculations indicate that the force generated per myosin per cycle was strongly frequency dependent [Fig. 3(b)].

III. DISCUSSION

Given the number of assumptions involved in the estimates used in the present calculations, some of which cannot be made with greater precision at the present time, the data of course have to be interpreted carefully. Changes in some assumptions would lead to greater calculated power values and changes in other assumptions would lead to smaller power values. The most important assumption, of course, is that in the ear canal at the level of the microphone the sound pressure and velocity are in phase. If this were not the case, the values we estimated would always be too small. The errors from this uncertainty can be very large and overshadow all the other possible errors accrued through incorrect assumptions. Since we cannot reconstruct the shape of the sound wavefront in the ear canal, the assumption of in-phase sound pressure and velocity is inescapable. We believe nonetheless that tentative conclusions can be reached.

A. The frequency map

In lizards, all papillae show a low-frequency area (up to about 1 kHz and apparently not involved in SOAE generation) and at least one high-frequency area that generates SOAE, sometimes in relatively large numbers (Manley, 1990; Köppl, 1995). Previous frequency-mapping data from five species (including *Gerrhonotus*, whose papilla is very similar to that of *Anolis*) shows that, despite a high degree of morphological variability, the frequency maps of different lizard species are remarkably uniform (Manley, 1990; Köppl and Manley, 1992). In these five species, the cells of the high-frequency area map the logarithm of frequency along the papilla's length, starting at about 1 kHz. Since the highest SOAE spectral peaks in *Anolis* were found at 7.7 kHz (at 30 °C), we assumed that the high-frequency end of the papilla codes for 8 kHz. This value is higher than previous estimates for other lizard species, including the very similar papilla of *Gerrhonotus* [4–4.3 kHz; see Manley (1990) for a review]. In *Gerrhonotus*, however, measurements were made at lower temperatures (21–24 °C) than we used, and both neural frequency tuning and SOAE frequency are known to be temperature sensitive (Klinke and Smolders, 1993; Manley and Köpple, 1994; Köppl, 1995). Correcting to 30 ° C, the upper limit for *Gerrhonotus* might exceed 6.5 kHz, depending on the temperature constant. In addition, we found that the stereovillar bundles of the hair-cell bundles at the high-frequency end of the papilla in *Anolis* are substantially shorter than those of *Gerrhonotus* (about 3 μm, instead of at least 7 μm; Mulroy and Williams, 1987). Shorter bundles are consistent with a higher upper frequency limit. Since the papillae of these two species are approximately the same length, this implies that the place-frequency gradient is substantially steeper in *Anolis* than in *Gerrhonotus*.

B. Hair-cell generators and myosin molecules

Whatever the generation mechanism underlying SOAE is, the present data demonstrate a correlation between emission power and the number of hair cells generating the emissions. This correlation supports the common assumption that hair cells are actually generating the SOAE. Our data thus add a further piece of evidence indicating that spontaneous hair-cell activity is the immediate cause of SOAE. The present data are consistent with assumptions made on other, less direct information, such as the detailed similarity between primary-neural tuning curves and SOAE suppression-tuning curves, and the hypoxia sensitivity of SOAE (e.g., Köppl and Manley, 1994).

Our data also demonstrate that it is possible that myosin molecules are in fact the motors generating SOAE, even at these relatively high frequencies. More recent refinements in the measurement of bundle motion have demonstrated the presence of fast reverse transduction with a sensitivity up to 0.6 nm/mV (Denk and Webb, 1992) and rapid twitching in hair-cell bundles of the frog sacculus (Benser *et al.*, 1996) in hair cells that are normally only sensitive to low frequencies. Since it has been demonstrated that evoked hair-cell *bundle* motility is associated with an increase in the sensitivity of transduction (Benser *et al.*, 1996), the bundle motility that we assume is generating the otoacoustic emissions may be intimately associated with the active process responsible for the high auditory sensitivity of nonmammalian hair cells. Support for the involvement of a type of myosin in hair-cell bundle motion can also be derived from the effects of solutions that influence contraction or relaxation in muscle systems on the motion of the bundles of hair cells from the Crista ampullaris (Orman and Flock, 1983).

According to the current understanding of mechanisms underlying outer hair-cell motility in mammals (see, e.g., Dallos and Evans, 1995), the above conclusions would suggest that mammalian and nonmammalian tetrapods have evolved different fundamental mechanisms underlying their active processes. Mammals may have at least partially abandoned an evolutionarily older mechanism (myosin-based active bundle motions) because of inherent limitations in the upper frequency limit.

Even in nonmammals, however, relatively high oscillation frequencies—in our lizard species up to 7.7 kHz—were observed. In another nonmammal, the barn owl, SOAE center frequencies up to 11 kHz have been measured (Manley and Taschenberger, 1993). Is it plausible that myosin systems could support such high oscillation rates? The fastest myosin-actin based muscle systems known are the stretch-activated, asynchronous flight muscles of insects, some of which show wing-beat frequencies above 1 kHz (e.g., in thrips; Tregear, 1975). Since the wing-beat frequency in such asynchronously controlled flight-muscle systems is determined by the mechanical resonance characteristics of the thorax and wings, removal of the wing loading raises the resonance frequency (Roeder, 1951; Satavalta, 1952). This rise in frequency can be up to 100% (Roeder, 1951) that, in the case of thrips, would mean thorax oscillation frequencies can rise above 2 kHz. Also, in some systems such as in bees, insects are known to ''sonicate'' (for example, during pollen

collection) using their flight muscles in a special mode in which the wings do not move and the oscillation frequency of the thorax is approximately double the normal wing-beat frequency (King *et al.*, 1996). Myosins operating in much smaller micromechanically tuned systems, such as hair-cell bundles in lizards, could presumably resonate at even higher frequencies, especially if it can be assumed that not every myosin molecule must contribute to every cycle of the oscillation.

How could the hair-cell bundle motors be stimulated into oscillation and entrained in frequency? Assuming that some internal or external noise source is responsible for initiating movement of the hair-cell bundle (e.g., ion-channel noise and membrane-potential noise; for a discussion, see Denk and Webb, 1992), stretch activation of the tip-link plaque myosins could initiate oscillation. The motors in any given hair cell would only react to stimulation in one direction, and the micromechanical resonance frequencies would then determine the frequency of the spontaneous oscillations. Micromechanical resonance has been directly shown in the alligator lizard and is also manifest in the frequency-selective tuning curves of hair cells and afferent nerve fibres in other lizard species (see, e.g., Manley, 1990; Köppl, 1995; Köppl and Manley, 1992). In other words, filtering of any kind of ''noise'' input to the bundles should suffice to activate the myosins, and these would always produce the strongest oscillation at the preferred frequency of the bundle. Frog sacculus hair cells with the otolith membrane removed show spontaneous bundle twitching (Benser *et al.*, 1996) that could form a basis for setting up spontaneous oscillations. Presumably then coupling between hair cells (through the fluid surrounding the bundles or through a tectorial membrane) could then lead to synchronous oscillations.

Except at lower frequencies, however, myosin activation is unlikely to be on a cycle-by-cycle basis for every single molecule. While in our data the calculated power per hair cell rises with frequency, as would be expected if the same motors are working at a higher frequency, the force per myosin falls toward higher frequencies (Fig. 3). This can be explained by assuming that the proportion of myosins that are active per oscillation cycle is falling as frequency rises. Thus the values for low frequencies, with a mean force at 1 kHz of near 0.1 pN, would be fairer estimates of the power produced per myosin molecule. This number is smaller than, but within an order of magnitude of the values given for myosins in vertebrate skeletal muscle at different temperatures (values range between 1 and 4 pN; Finer *et al.*, 1994; Huxley and Simmons, 1971; Ishijima *et al.*, 1991). However, since we know neither whether the number of myosins per insertional plaque is constant in all hair cells along the epithelium nor whether all myosins have the same stroke length and produce the same power per stroke, further speculation is not yet timely.

Note added in proof: Magnan *et al.* (1997), measuring the reverse transfer function of the middle ear in the guinea pig, recently reported that distortion products suffer a loss in amplitute of about 30 dB between the cochlea and the ear canal. This loss is relatively flat across frequency. If the same is true in our lizard species, then the mean force generated by each myosin at 1 kHz would be about 1 pN, very close to the values given for vertebrate skeletal muscle (1 to 4 pN).

An additional problem associated with calculation of the power produced by myosins is the assumption that the force produced in the bundle couples into the inner-ear fluid system and out into the ear canal with the same efficiency at all frequencies. Given that the hair-cell bundles change their height by up to seven times along the relevant area of the papilla, this assumption is also highly unlikely to be correct. In other species, where the hair cells are connected to a tectorial membrane, the coupling through the tectorial membrane may be sufficient to entrain all local hair cells to the same phase of oscillation. In such cases, also, a coordinated spontaneous oscillation could crystallize out of noisy movements in coupled cells. We have reported that SOAE peaks are, indeed, larger and fewer in number in lizard papillae in which the hair cells are more strongly coupled through the tectorial membrane (Manley, 1996). Thus in spite of the above caveats, the working hypothesis we have adopted is that the adaptational myosin motors are indeed capable of producing the energy manifest in lizard SOAE.

An interesting additional speculation follows from the finding of Benser *et al.* (1996) that rapid spontaneous twitches of bundles of hair cells from the bullfrog's sacculus were each associated with transient membrane depolarizations in the hair cell. If these depolarizations lead to a transient increase in transmitter release at the afferent synapse, then spontaneous mechanical twitching in the bundle could lead to spontaneous transmitter release. If, as we believe, the twitching is integrated into a spontaneous oscillation of the bundle at or near the characteristic frequency of the hair cell, the intervals between the pulses of transmitter release would be the reciprocal of the characteristic frequency. Precisely this kind of ''preferred interval'' has been observed in time-interval histograms of the spontaneous discharge activity of afferent fibers of nonmammals, but not in mammals (summarized in Manley, 1990).

ACKNOWLEDGMENTS

We thank H. Fastl and S. Wartini for help with probe optimization and with the calculations of acoustic power, C. Köppl and A. Pavusa for assistance with the anatomy, and J. Ashmore, C. D. Geisler, O. Gleich, C. Köppl, J. O. Pickles, and G. K. Yates, who commented on earlier versions of the manuscript. Supported by the Deutsche Forschungsgemeinschaft within the program of the SFB 204 (Gehör).

Alexander, R. M. (**1995**). ''Springs for wings,'' Science **268,** 50–51.

Assad, J. A., and Corey, D. P. (**1992**). ''An active motor model for adaptation by vertebrate hair cells,'' J. Neurosci. **12,** 3291–3309.

Assad, J. A., Shepherd, G. M. G., and Corey, D. P. (**1991**). ''Tip-link integrity and mechanical transduction in vertebrate hair cells,'' Neuron **7**, 985–994.

Benser, M. E., Marquis, R. E., and Hudspeth, A. J. (**1996**). ''Rapid, active hair-bundle movements in hair cells from the bullfrog's sacculus,'' J. Neurosci. **16,** 5629–5643.

Crawford, A. C., and Fettiplace, R. (**1985**). ''The mechanical properties of ciliary bundles of turtle cochlear hair cells,'' J. Physiol. (London) **364,** 359–379.

Dallos, P., and Evans, B. N. (**1995**). ''High-frequency motility of outer hair cells and the cochlear amplifier,'' Science **267,** 2006–2009.

Denk, W., and Webb, W. W. (**1992**). ''Forward and reverse transduction at the limit of sensitivity studied by correlating electrical and mechanical fluctuations in frog saccular hair cells,'' Hearing Res. **60,** 89–102.
Finer, J. T., Simmons, R. M., and Spudich, J. A. (**1994**). ''Single myosin molecule mechanics: piconewton forces and nanometre steps,'' Nature (London) **368,** 113.
Gillespie, P. G., Wagner, M. C., and Hudspeth, A. J. (**1993**). ''Identification of a 120 kd hair-bundle myosin located near stereociliary tips,'' Neuron **11**, 581–594.
Howard, J., and Hudspeth, A. J. (**1988**). ''Compliance of the hair bundle associated with gating of mechanoelectrical channels in the bullfrog's saccular hair cells,'' Neuron **1**, 189–199.
Howard, J., Roberts, W. M., and Hudspeth, A. J. (**1988**). ''Mechanoelectrical transduction by hair cells,'' Ann. Rev. Biophys. Biophys. Chem. **17**,
Hudspeth, A. J., and Gillespie, P. G. (**1994**). ''Pulling springs to tune transduction: Adaptation by hair cells,'' Neuron **12**, 1–9.
Huxley, H. E. (**1990**). ''Sliding filaments and molecular motile systems,'' J. Biol. Chem. **265**, 8347–8350.
Huxley, A. F., and Simmons, R. M. (**1971**). ''Proposed mechanism of force generation in striated muscle,'' Nature (London) **233,** 533–538.
Ishijima, A., Doi, T., Sakurada, K., and Yanagida, T. (**1991**). ''Subpiconewton force fluctuations of actomyosin *in vitro*,'' Nature (London) **352,** 301–306.
King, M. J., Buchmann, S. L., and Spangler, H. (**1996**). ''Activity of asynchronous flight muscle from two bee families during sonication (Buzzing),'' J. Exp. Biol. **199**, 2317–2321.
Klinke, R., and Smolders, J. W. Th. (**1993**). ''Performance of the avian inner ear,'' Prog. Brain Res. **97,** 31–43.
Köppl, C. (**1995**). ''Otoacoustic emissions as an indicator for active cochlear mechanics: A primitive property of vertebrate hearing organs,'' in *Advances in Hearing Research*, edited by G. A. Manley, G. M. Klump, C. Köppl, H. Fastl, and H. Oeckinghaus (World Scientific, Singapore), pp. 207–218.
Köppl, C., and Manley, G. A. (**1992**). ''Functional consequences of morphological trends in the evolution of lizard hearing organs,'' in *The Evolutionary Biology of Hearing*, edited by R. R. Fay, A. N. Popper, and D. B. Webster (Springer-Verlag, New York), pp. 489–509.
Köppl, C., and Manley, G. A. (**1994**). ''Spontaneous otoacoustic emissions in the bobtail lizard. II: Interactions with external tones,'' Hearing Res. **72,** 159–170.
Magnan, P., Avan, P., Dancer, A., Smurzynski, J., and Probst, R. (**1997**). ''Reverse middle-ear transfer function in the guinea pig measured with cubic difference tones,'' Hearing Res. **107,** 41–45.
Magnan, P., Probst, R., Smuzynski, J., Avan, P., and Dancer, A. (**1996**). ''Direct measurement of the reverse transfer function of the guinea pig middle ear,'' in *Diversity in Auditory Mechanics*, edited by E. Lewis, G. Long, P. Narins, and C. Steele (World Scientific, Singapore) (in press).
Manley, G. A. (**1990**). *Peripheral Hearing Mechanisms in Reptiles and Birds* (Springer-Verlag, Heidelberg).
Manley, G. A. (**1995**). ''The avian hearing organ: A status report,'' in *Advances in Hearing Research*, edited by G. A. Manley, G. M. Klump, C. Köppl, H. Fastl, and H. Oeckinghaus (World Scientific, Singapore), pp. 219–229.
Manley, G. A. (**1996**). ''Diversity in hearing-organ structure and the characteristics of spontaneous otoacoustic emissions in lizards,'' in *Diversity in Auditory Mechanics*, edited by E. R. Lewis, G. R. Long, R. F. Lyon, P. M. Narins, and C. R. Steele (World Scientific, Singapore) (in press).
Manley, G. A., and Johnstone, B. M. (**1974**). ''Middle-ear function in the guinea pig,'' J. Acoust. Soc. Am. **56,** 571–576.
Manley, G. A., and Köppl, C. (**1994**). ''Spontaneous otoacoustic emissions in the bobtail lizard III: Temperature effects,'' Hearing Res. **72,** 171–180.
Manley, G. A., and Taschenberger, G.(**1993**). ''Spontaneous otoacoustic emissions from a bird: A preliminary report,'' in *Biophysics of Hair Cell Sensory Systems*, edited by H. Duifhuis, J. Horst, P. van Dijk, and S. van Netten (World Scientific, Singapore), pp. 33–39.
Manley, G. A., Gallo, L., and Fastl, H. (**1995**). ''Spontaneous otoacoustic emissions, tip links, and myosin.'' Abstract No. **372**, 18th Midwinter meeting, Assoc. Res. Otolaryngol.
Miller, M. R. (**1992**). ''The evolutionary implications of the structural variations in the auditory papilla of lizards,'' in *The Evolutionary Biology of Hearing*, edited by R. R. Fay, A. N. Popper, and D. B. Webster (Springer-Verlag, New York), pp. 463–487.
Mulroy, M. J., and Williams, R. S. (**1987**). ''Auditory stereocilia in the alligator lizard,'' Hearing Res. **25,** 11–21.
Orman, S., and Flock, Å. (**1983**). ''Active control of sensory hair mechanics implied by susceptibility to media that induce contraction in muscle,'' Hearing Res. **11,** 261–265.
Pickles, J. O., Comis, S. D., and Osborne, M. P. (**1984**). ''Cross-links between stereocilia in the guinea pig organ of Corti and their possible relation to sensory transduction,'' Hearing Res. **15,** 103–112.
Roeder, K. (**1951**). ''Movements of the thorax and potential changes in the thoracic muscles of insects during flight,'' Biol. Bull. Mar. Biol. Lab., Woods Hole **100**, 95–106.
Satavalta, O. (**1952**). ''Wing-stroke frequency of insects,'' Ann. Zool.-Botan. Fenn. **15**, 1–47.
Sparrow, J. C. (**1995**). ''Flight and phosphorylation,'' Nature (London) **374,** 592–593.
Tregear, R. T. (**1975**). ''The biophysics of fibrillar flight muscle,'' in *Insect Muscle*, edited by P. N. R. Usherwood (Academic, New York), pp. 357–403.
Wever, E. G. (**1978**). *The Reptile Ear* (Princeton U. P., Princeton, NJ).
Zizz, C. A., and Glattke, T. J. (**1988**). ''Reliability of spontaneous otoacoustic emission suppression tuning curve measures,'' J. Speech Hear. Res. **31,** 616–619.
Zwicker, E., and Fastl, H. (**1990**). *Psychoacoustics—Facts and Models* (Springer-Verlag, Heidelberg).

The representation of concurrent vowels in the cat anesthetized ventral cochlear nucleus: Evidence for a periodicity-tagged spectral representation

Suzanne E. Keilson[a)]
Center for Hearing Sciences and Department of Biomedical Engineering, Johns Hopkins University, Baltimore, Maryland 21205

Virginia M. Richards
Department of Psychology, University of Pennsylvania, Philadelphia, Pennsylvania 19104

Bradley T. Wyman and Eric D. Young
Center for Hearing Sciences and Department of Biomedical Engineering, Johns Hopkins University, Baltimore, Maryland 21205

(Received 8 August 1996; revised 22 January 1997; accepted 23 January 1997)

Chopper units of the ventral cochlear nucleus (VCN) provide a rate representation of stimulus spectrum and a temporal representation of fundamental frequency (F_0). This dual representation may be useful in segregating competing speech sounds, where differences in F_0 are a cue. Responses to the vowel portion of concurrently presented pairs of syllables /bV∫/ with different F_0's (88, 98, and 112 Hz) were studied in the VCN of anesthetized cats; 11 English vowels were used for V. Vowels were chosen so that one had a formant frequency just above the unit's best frequency (BF) and the other had a formant just below BF. By changing the stimulus sampling rate, formant peaks were shifted relative to the unit's BF, producing a range of stimuli, varying in the relative power of the two vowels within the unit's tuning curve. Results show that units' discharge rates reflect the energy within their tuning curves and the relative synchronization of units' responses to the two F_0's favors the dominant vowel. A method of segregating two vowels is provided in which relative synchronization to the F_0's is used to apportion discharge rate between the vowels. Best results were obtained in chopper units, although primarylike units showed similar behavior. [S0001-4966(97)00906-5]

PACS numbers: 43.64.Qh, 43.64.Sj, 43.71.Cq [RDF]

INTRODUCTION

Separation of a target speech stream from a noisy background is a difficult process when the background is competing speech because the spectral similarity of the target and the interfering sound makes filtering of little use. Several cues are available which make speech more intelligible when masked by noise or other speech, including separation of the sources in space (Bronkhurst and Plomp, 1988), grouping cues such as simultaneity in the onset and offset of components at different frequencies (Darwin, 1981, 1984), and differences in fundamental frequency (F_0) between the target and interfering speech (Brokx and Nooteboom, 1982). This paper concerns a possible neural correlate of the last cue, differences in F_0.

An important stimulus for the study of separation of simultaneous speech streams has been the double vowel, consisting of two vowels presented simultaneously to the same ear. This stimulus eliminates most of the known cues and allows the effects of differences in F_0 to be studied. Performance in identifying both of the vowels of a double-vowel stimulus improves as the F_0's of the vowels are separated (Scheffers, 1983; Assmann and Summerfield, 1990, 1994; Culling and Darwin, 1993). The improvement is rapid as the difference in F_0 grows from zero to one semitone, and is slower for larger differences. Differences in F_0 are not the only cue even with this stimulus, however, since performance is well above chance with no difference in F_0 (Assmann and Summerfield, 1989). Despite this fact and despite the fact that performance may not improve with differences in F_0 in all conditions (such as with short-duration stimuli; see Assmann and Summerfield, 1990), the results strongly suggest that differences in F_0 between vowels aid in their separation and identification.

Models of speech separation in the double-vowel paradigm and other situations have usually assumed that differences in F_0 aid in vowel separation by identifying the vowel to which components of a double-vowel stimulus belong (Parsons, 1976; Weintraub, 1987; Summerfield, 1988; Assmann and Summerfield, 1990; Meddis and Hewitt, 1992; de Cheveigné, 1993; de Cheveigné *et al.*, 1995). These models consist of two stages: (1) a stage which estimates the F_0's of one or both vowels; and (2) a stage which uses the identified F_0's to separate the vowels, usually by selecting or canceling the components of the stimulus which are harmonics of one fundamental. Perceptual evidence provides only partial support for this general class of model. The use of composite vowels in which the F_0's of the formants differ has allowed analysis of the strength of F_0 as a cue for grouping components together, i.e., the second stage in the models above (Culling and Darwin, 1993, 1994); these studies suggest that F_0 is a grouping cue only for larger differences in

a)Electronic mail: keilson@loyola.edu

F_0 ($>2-4$ semitones). At smaller differences, the improvement in performance with F_0 difference is based on as-yet poorly understood waveform interactions, such as beats between the harmonics of the two vowels (Assmann and Summerfield, 1994; Culling and Darwin, 1994); these waveform cues are enhanced by small differences in F_0, although in different ways than is assumed in the models discussed above.

In the peripheral auditory system, the neural representation of speech is assumed to be tonotopic and therefore has been analyzed by measuring and plotting the neural response versus best frequency (BF; reviewed by Sachs, 1984); that is, each neuron is assumed to report on the energy in the stimulus at frequencies near its BF. There are two classes of representation, depending on the definition of neural response. The simplest response measure is average discharge rate during the stimulus, the so-called rate-place measure. At the level of the auditory nerve (AN), the rate-place measure provides sufficient information to allow single vowels to be identified or discriminated (Sachs *et al.*, 1988; Conley and Keilson, 1995; May *et al.*, 1996); however, the rate-place measure is nonrobust in that it changes dramatically with signal level or background noise masking (Sachs and Young, 1979; Sachs *et al.*, 1983; Sinex and Geisler, 1983; Delgutte and Kiang, 1984c, d). Moreover, rate-place measures have been shown not to provide sufficient spectral detail to support perceptual changes in vowel identity produced by changing the phase of a single harmonic (Palmer *et al.*, 1987). The second kind of response measure is based on phase locking or the temporal patterns of AN spike trains; several measures of this type have been defined (Young and Sachs, 1979; Sachs and Young, 1980; Sinex and Geisler, 1983; Delgutte and Kiang, 1984a, b). These measures are much more robust than the rate-place measure and have been shown to provide an adequate basis for F_0-based double-vowel segregation (Palmer, 1990; de Cheveigné, 1993).

In the ventral cochlear nucleus (VCN) auditory information is carried in a number of parallel pathways (reviewed by Cant, 1992; Rhode and Greenberg, 1992). The work described here focuses on two of these, the primarylike and chopper units. Primarylike units resemble AN fibers in many of their properties (Bourk, 1976; Blackburn and Sachs, 1989) including their representation of vowel stimuli (Palmer *et al.*, 1986; Blackburn and Sachs, 1990; Winter and Palmer, 1990b). By contrast with AN fibers, choppers provide an enhanced rate-place representation, which contains good information about the spectrum of a single-vowel stimulus (Blackburn and Sachs, 1990); the representation is robust against changes in stimulus level. At the same time, choppers provide little in the way of a temporal representation, since they phase lock to the vowel waveform only up to $\approx 500-700$ Hz, i.e., through the first formant region at best (Bourk, 1976; Blackburn and Sachs, 1989; Winter and Palmer, 1990a). At all BFs, choppers' temporal responses to single vowels contain a significant component at the fundamental frequency of the vowel; at BFs above 700 Hz, this component dominates the phase locking (Blackburn and Sachs, 1990). Thus choppers potentially provide a *periodicity-tagged spectral representation*, in which average discharge rate encodes for stimulus energy and the temporal pattern of response identifies the F_0 of the stimulus (loosely called periodicity tagging) which provides that energy. Note that this sort of periodicity tagging is similar to the operation of the harmonic-segregation models mentioned above.

This possibility of a periodicity-tagged spectral representation in VCN choppers is the hypothesis that was examined in this study. The focus of the study was on double-vowel stimuli with different F_0's and on whether the properties described above for single vowels would hold for this more complex stimulus. Because of the difficulty of obtaining a large population of units in VCN, the analysis was done by a pseudo-population approach in which a single unit is used to approximate a population of units with various BFs by shifting the stimulus along the frequency axis (instead of studying a large group of units with different BFs using the same stimulus). When this shifting is done, a chopper's rate responses follow the profile of stimulus energy near the unit's BFs and its phase locking is dominated by the F_0 of the vowel with the larger amount of energy within its tuning curve. This response pattern supports the idea of a periodicity-tagged spectral representation in choppers. The same sort of representation will also be shown for primary-like units, although the quality of the representation is not as good.

I. METHODS

A. Preparation, acoustics, and unit recording

Nine domestic short-hair adult cats were used in the experiments; otoscopic examination showed the cats to be free of middle-ear infections and external-ear obstructions. The animals were premedicated with atropine (0.2 cc im), tranquilized with xylazine (0.1 cc im), and anesthetized with ketamine (1 cc im). During the surgery and experiment, anesthesia was maintained with IV injections of sodium pentobarbital (≈ 5.5 mg/h) to maintain areflexia. A tracheal cannula was inserted; the bulla was exposed and vented with ≈ 30 cm of PE-90 tubing. The cat was placed in an electrically shielded soundproof chamber (IAC-1204A) on a regulated heating pad which maintained body temperature at 36–38 °C. The skull was opened from the nuccal ridge to the bony tentorium and the VCN was exposed by cerebellar aspiration. Single unit recordings were done with platinum-black-tipped Pt-Ir microelectrodes placed in the VCN under visual control and advanced by a hydraulic microdrive. Histological controls were not done, but recordings were certainly made in the VCN, based on the known properties units in VCN versus those in dorsal CN (Rhode and Greenberg, 1992).

Acoustic stimuli were delivered to the ipsilateral ear through a closed acoustic system connected to a hollow earbar (Sokolich, 1977). Prior to data collection, a calibration of the sound delivery system was performed in the cat's ear using a probe tube placed at the mouth of the earbar. The calibration functions were reasonably flat (± 7 dB) in the

TABLE I. Parameters used to synthesize the CVC syllables. Stimuli take the form /bV∫/, where V is one of 11 vowels (left two columns); stimuli were generated with three fundamental frequencies (F_0 3rd column; exact F_0's are 87.84, 97.60, and 112.24 Hz). Formant parameters were taken from Peterson and Barney (1952). The first three formants were adjusted slightly to lie on a harmonic of F_0. For all vowels, $F4=3.3$ kHz (BW4=0.25 kHz) and $F5=3.75$ kHz (BW5=0.2 kHz). The syllable-leading /b/ was synthesized by starting $F1$, $F2$, and $F3$ at 0.2, 1.1, and 2.15 kHz, respectively; the formant transitions lasted 40.96 ms and were linearly interpolated from the starting frequencies to the vowel formants. /b/ and the vowel were synthesized with five cascade resonators. The syllable-terminating /∫/ was produced with 40.96-ms formant transitions to 0.2, 2.4, and 4.8 kHz, followed by broadband noise driving parallel resonators at those three formant frequencies.

Vowel	Example	F_0 (Hz)	$F1$ (kHz)	$F2$ (kHz)	$F3$ (kHz)	$BW1$ (kHz)	$BW2$ (kHz)	$BW3$ (kHz)
i	beet	88	0.264	2.285	2.989	0.05	0.20	0.40
		98	0.293	2.247	3.029			
		112	0.224	2.240	3.024			
I	bit	88	0.352	2.022	2.549	0.05	0.10	0.14
		98	0.391	1.954	2.540			
		112	0.336	2.016	2.576			
ɛ	bet	88	0.527	1.846	2.461	0.06	0.09	0.20
		98	0.489	1.856	2.443			
		112	0.560	1.792	2.464			
æ	bat	88	0.703	1.758	2.373	0.07	0.15	0.32
		98	0.684	1.759	2.443			
		112	0.672	1.680	2.464			
ʌ1	$but_{modif.}$	88	0.791	1.494	2.989	0.09	0.10	0.20
		98	0.782	1.466	3.029			
		112	0.784	1.456	3.024			
ɝ	bird	88	0.527	1.319	1.670	0.10	0.06	0.11
		98	0.489	1.368	1.661			
		112	0.448	1.344	1.680			
ʌ2	but	88	0.615	1.231	2.373	0.08	0.05	0.14
		98	0.684	1.172	2.345			
		112	0.672	1.232	2.352			
a	hot	88	0.703	1.055	2.461	0.13	0.07	0.16
		98	0.684	1.075	2.443			
		112	0.784	1.120	2.464			
u	boot	88	0.264	0.879	2.198	0.065	0.11	0.14
		98	0.293	0.879	2.247			
		112	0.336	0.896	2.240			
U	foot	88	0.440	1.055	2.198	0.08	0.10	0.08
		98	0.489	0.977	2.247			
		112	0.448	1.008	2.240			
ɔ	baud	88	0.527	0.879	2.373	0.09	0.10	0.08
		98	0.586	0.879	2.443			
		112	0.560	0.896	2.464			

frequency range of interest (<5 kHz); no corrections for irregularities in the calibration were applied to the acoustic stimuli.

B. Stimuli

The goal of the experiments was to study the responses of units to simultaneously presented pairs of vowels, with the vowels of a pair having different fundamental and formant frequencies. The major variable of interest is the relative energy in the two vowels at frequencies near the unit's BF. In order to accomplish this aim, we synthesized 11 stimuli, each containing a different English vowel; the formant frequencies of the vowels were chosen to cover the frequency range below 3 kHz as uniformly as possible, so that it would always be possible to choose two vowels with formants near a unit's BF. The stimuli were consonant-vowel-consonant (CVC) syllables of the form /bV∫/, where V is a vowel (e.g., ''bish,'' ''besh,'' ''bash,'' etc.). Although the stimuli contained consonants, and therefore frequency transitions and changes in source characteristics, a portion of each stimulus contained a periodic vowel with fixed formant frequencies; in this paper, data are reported only for responses during this vocalic portion.

Syllables containing 11 different vowels were generated using 3 fundamental frequencies ($F_0=87.84$, 97.60, and 112.24 Hz, called 88, 98 and 112 below); the differences between fundamental frequencies in terms of semitones are 1.82 (88 vs 98), 2.42 (98 vs 112) and 4.24 (88 vs 112). F_0 was constant throughout each stimulus. Table I shows the formant frequencies and bandwidths of the stimuli; the characteristics of the two consonants and other temporal properties of the CVC are given in the caption of Table I. The first three formant frequencies of each stimulus were adjusted slightly to lie exactly on harmonics of F_0; the numbers in Table I do not appear consistent with this statement because

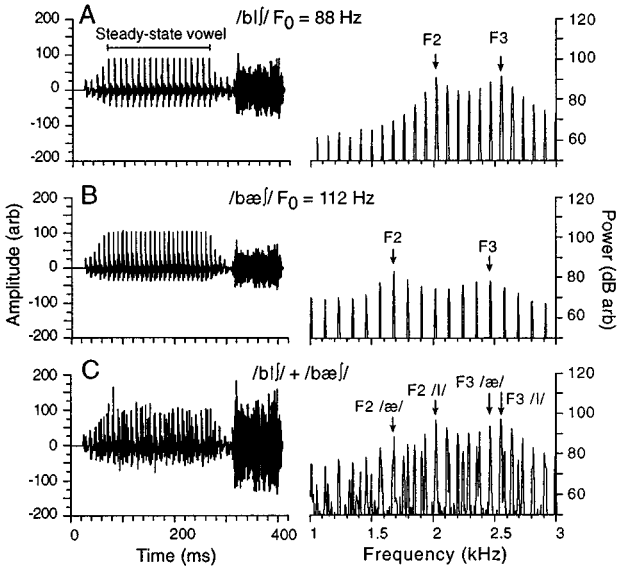


FIG. 1. Time-domain waveforms (left) and a portion of the spectra (right) of two syllables and their sum. The line at the top of the time-domain plot in (A) shows the portion of the stimulus where the formants are constant at their vocalic values. Note that the spectra are shown over a limited range of frequencies near the second and third formants ($F2$ and $F3$). (A) The syllable /bɪʃ/ with F_0 88 Hz. (B) The syllable /bæʃ/ with F_0 112 Hz. (C) The sum of the two stimuli in (A) and (B).

of roundoff error. The stimuli were synthesized with a Klatt synthesizer (Klatt, 1980) at a sampling rate of 50 kHz. The synthesizer was implemented in the ACSL simulation language and used both cascaded and parallel resonators to model the vocal tract transfer function. The stimuli were filtered with a transfer function equal to the human external ear transfer function (Wiener and Ross, 1946); the net effect was that the stimuli have roughly the same spectral shape at the cat eardrum that they would have at a human eardrum. The entire syllable lasted 409.6 ms while the steady-state vowel portion was 204.8 ms, when played at the synthesis sampling rate of 50 kHz. In the text and the figures, vowels are identified by their symbol and their approximate fundamental frequency, e.g., as /I/ (88 Hz).

The 33 stimuli were saved on disk and presented singly or in pairs through two D/A converters. When two stimuli were presented simultaneously, the analog signals were attenuated individually and then added together at the input of a second attenuator. Figure 1 shows an example of two stimuli and their sum. The left column shows the time-domain stimuli and the right column shows an expanded portion of the spectra of the vocalic part of the syllables, centered on the second and third formant regions. The line at the top of the left part of Fig. 1A marks the periodic vocalic portion of the stimulus. Figure 1A show the stimulus /bIʃ/ and Fig. 1B shows /bæʃ/, while Fig. 1C shows their sum. The second and third formants of the two vowels are marked on the spectrum plots. These two vowels might have been chosen for a unit with a BF near 1.8 kHz, between the second formants of the two vowels. The stimulus situation shown in Fig. 1 is typical for our data, because most units studied had BFs near the second or third formants.

C. Protocol

Single units were isolated using broadband noise as a search stimulus. Once a unit was isolated, its tuning curve (TC) was measured using an automated program and the best frequency (BF) was determined. Two types of measurements were taken to allow classification of unit type: (1) rate-level functions for noise and for BF tones were determined by presenting stimuli in 1-dB steps (one repetition of each) over a 100-dB range; and (2) PST histograms (0.1-ms bins) were computed from responses to 50-ms BF tones at 20, and 30 dB *re*: threshold. Units were classified according to first-spike latency and PST histogram shape as chopper, primarylike (pri), or primarylike-with-notch (pri-N) using standard criteria (Bourk, 1976; Young *et al.*, 1988; Blackburn and Sachs, 1989). Choppers were further subdivided into chop-S and chop-T categories based on regularity of discharge (Young *et al.*, 1988; Blackburn and Sachs, 1989). Onset units, which are the remaining major response class in VCN, were not studied. In the text below, the term ''primarylike'' will be used to refer to the pri and pri-N subtypes collectively. For the details of these classification methods, the summary in Blackburn and Sachs (1989) should be consulted.

After a unit was isolated and characterized, two syllables were selected from the list in Table I such that one vowel had a formant (usually the second) just above the unit's BF and the other vowel had a formant just below the unit's BF. The two syllables also had different F_0's. In order to determine suitable levels at which to present the stimuli, rate-level functions for each syllable by itself were measured and the unit's thresholds for each syllable were determined. The double-syllable stimuli were then presented with both syllables at 20 dB above threshold and, time allowing, repeated at 30 dB above threshold. These levels ensured that the vowels were presented within the dynamic range of the unit. Each stimulus was presented 100 times at a 1-s interval. Stimulus presentations were repeated with the spectra of the stimuli shifted along the log-frequency axis (see next section) by various amounts. The shifts were chosen to place one vowel's formant at the unit's BF, then the other vowel's formant at BF, then to place the unit's BF approximately halfway between the formants. If time permitted, other shifts were also used, so that data were available with the unit's BF at several different locations with respect to the formants of the two stimuli.

D. Frequency shifting of stimuli

Shifting the stimulus spectrum was done by changing the sampling rates of the D/A converters. The two D/A converters, one for each vowel, had the same sampling rate, so increasing (or decreasing) the D/A sampling rate had the effect of shifting the stimulus spectrum up (or down) on a log frequency scale, without changing the ratios of the formants of the two stimuli. The shift is illustrated in Fig. 2A–C which shows the spectrum of a vowel when played back at twice the synthesis rate (A), at the synthesis rate (B, 50 kHz), and at half the synthesis rate (C). A unit's tuning curve is plotted on the same frequency axis to illustrate the changing

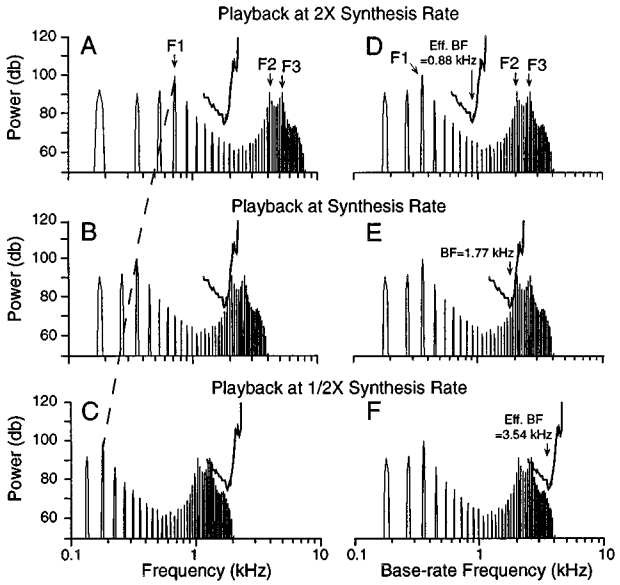


FIG. 2. The relationship between a unit's tuning curve (BF=1.77 kHz) and a single-vowel spectrum (/bɪʃ/) at three playback sampling rates F_P: 100 kHz in (A) and (D); 50 kHz (equal to the synthesis rate) in (B) and (E); and 25 kHz in (C) and (F). In the left column, plots are shown on the true frequency axis. In the right column, plots are shown on the base-rate frequency axis b_k, in which the tuning curves are shifted and the spectrum is not; see text and Eq. (4).

relationship between the unit's BF and the features of the stimulus. Note that changes in playback rate also change stimulus duration.

The plots in Fig. 2A–C shows the stimulus and the tuning curve on the true frequency axis. For the analysis of the data, we consider the frequency axis as it is shown in Fig. 2D–F, in which the stimulus spectrum remains constant and the unit's BF changes. To assist in describing this method of presenting the data we introduce two new terms. The frequency axis of Fig. 2D–F is called the *base-rate frequency*, since the stimulus components stay at the same frequencies as if the stimulus were played back at the original synthesis sampling rate (50 kHz). The base-rate frequency of a particular feature of the stimulus (e.g., a formant) is independent of the actual playback rate and is related to the actual frequency of the feature as

$$\text{Actual frequency} = \frac{\text{Playback sampling rate}}{\text{Synthesis sampling rate}} \cdot \text{Base-rate frequency}. \quad (1)$$

The *effective BF* is the BF of the unit when the spectrum is plotted in terms of base-rate frequency, i.e.,

$$\text{Effective BF} = \frac{\text{Synthesis sampling rate}}{\text{Playback sampling rate}} \cdot \text{Actual BF}. \quad (2)$$

The advantage of plotting our results in terms of base-rate frequency and effective BF is that stimulus features, including F_0's and formants, remain fixed, while their relationship to the tuning curve of the unit varies.

Comparison of the two columns of Fig. 2 shows that the two ways of looking at frequency preserve the relationship between the tuning curve and stimulus spectrum, on a log-frequency axis. As the stimulus spectrum is shifted, the responses of a unit provide an approximation of the responses to one stimulus from units with different BFs, to the extent that the log-frequency scale approximates the layout of frequencies in the cochlea. The approximation is inexact for three reasons: first, the log-frequency approximation breaks down at low frequencies (below about 1–2 kHz in cat, Keithley and Schreiber, 1987). Second, tuning curves do not show a constant bandwidth/BF ratio across BF (Liberman, 1984). Third; there are changes with BF in such properties as two-tone suppression (Prijs, 1989) and strength of phase locking (Johnson, 1980; Bourk, 1976). In this paper, we use the frequency-shift approximation over relatively narrow frequency ranges, so these effects should not invalidate the results.

As described above, playback sampling rates were selected so that the effective BF would: (1) fall between the frequency F_{V1} of a formant of one vowel and the frequency F_{V2} of a formant of the second vowel (e.g., the second formants of the stimuli in Fig. 1, for a 1.8-kHz unit); (2) fall on F_{V1}; (3) fall on F_{V2}; (4) fall on the geometric mean of cases 1 and 2; (5) fall on the geometric mean of cases 1 and 3. Usually the vowels were chosen so that case 1 occurred when the playback rate equaled the synthesis rate. In later studies the strategy was expanded to move the effective BF beyond the formants and add additional points between the formants.

E. Response measures

For each set of responses to 100 repetitions of a stimulus, the following response measures were computed:

(1) Average discharge rate during the vowel portion of the stimulus was computed. Spikes for the rate calculation were taken from a 102.4-ms window beginning 112.64 ms into the stimulus. These times apply when the playback sampling rate equaled the synthesis sampling rate and were adjusted for the change in stimulus duration when the playback sampling rate was changed.
(2) A post-stimulus-time histogram (PSTH) of 4096 bins was generated using a binwidth BW of 50 μs (at the playback sampling rate); BW varied inversely with playback sampling rate so that the stimulus occupied a constant fraction of the PSTH abscissa.
(3) The discrete Fourier transform (DFT) of the PSTH was computed after a 204.8-ms wide Kaiser window was applied to the steady-state vowel portion, beginning 61.44 ms into the stimulus (again, these numbers were adjusted for the playback sampling rate); the DFT values were converted to synchronization index by dividing each bin in the DFT by the zero-frequency bin (which is the mean discharge rate if the PSTH is expressed in units of rate).

That is, if $p(n)$ is the PSTH, $P(k)$ is the complex DFT of the PSTH, and $S(k)$ is the synchronization index, then

$$P(k)=\sum_{n=0}^{N-1} p(n)\exp(-j2\pi kn/N),$$
$$S(k)=\frac{|P(k)|}{|P(0)|}, \tag{3}$$

where N is the number of bins in the PSTH, and k and n are discrete indices. PSTHs are plotted below in terms of real time, equal to $N\cdot\mathrm{BW}$ for the nth bin; the true frequency of DFT component $P(k)$ is $f_k=k/(N\cdot\mathrm{BW})$. DFTs are plotted in terms of base-rate frequency b_k where

$$b_k=f_k\frac{50\,000}{F_P}=\frac{k}{N\cdot\mathrm{BW}}\frac{50\,000}{F_P} \tag{4}$$

and F_P is the playback sampling frequency in Hz.

II. RESULTS: TEMPORAL RESPONSES TO DOUBLE-VOWEL STIMULI

A total of 64 data sets were obtained from 37 units. A data set is a complete experiment consisting of responses to one pair of syllables at five or more values of F_p. Some units with stable isolation were tested with either different vowels or different stimulus levels. Ten of these units (19 data sets) were discarded because they could not be classified as primarylike or chopper or because of poor isolation, usually caused by large neurophonics (typical of low BF regions of the CN). Seven units (16 data sets) were chop–S, 9 units (13 data sets) were chop-T, 4 units (4 data sets) were pri, and 7 units (12 data sets) were pri-N. Most of these data were obtained using the second (15 units, 25 data sets) or third (10 units, 16 data sets) formants; data were obtained using the first formant in 2 units (4 data sets).

A. Chopper units

A typical example of the responses of a chop-T unit to two vowels presented alone and concurrently is shown in Fig. 3. The unit's BF is 1.77 kHz and the vowels have second formants above (/I/$F2=2.022$ kHz) and just below (|æ| $F2=1.680$ kHz) the BF. The first column plots the PSTH for the steady-state vowel portion of the syllables and the second and third columns plot the magnitude of the discrete Fourier transform of the PSTH over two different frequency ranges. The abscissae of the Fourier transforms are scaled in terms of base-rate frequency so that when F_p changes, vowel features remain at the same location. However, the effective BF, which is marked in the right column, moves relative to the stimulus when F_P changes. The vowels' formant frequencies are marked on the plots in the right column; the dots in the center column indicate the F_0's of the vowels.

The top row shows the response to the vowel /I/ with $F_0=88$ Hz. As has been shown previously for choppers (Blackburn and Sachs, 1989), there is little or no synchronized response to the second formant of the vowel, even though it is near the unit's BF (right column). Instead, the response is dominated by the fundamental and its harmonics up to ≈ 1 kHz; this response is clear in the Fourier transform in the center column. The bottom row shows the response to the steady-state portion of the second vowel /æ/ with $F_0=112$ Hz. Again, there is a strong synchronization to F_0 and a small synchronization near the second formant of /æ/.

The second through sixth rows of Fig. 3 show responses when both vowels are presented together. Running from the second to the sixth row, the sampling rate, F_P, is systematically increased from 41.2 kHz to 53.6 kHz. As can be seen from the arrows in the right column, this moves the effective BF from a position slightly above the second formant of /I/ (88 Hz) to a position slightly below the second formant of /æ/ (112 Hz). Notice that the F_0 synchronization changes from predominantly locked to base-rate 88 Hz to predominantly locked to base-rate 112 Hz as the sampling rate is increased (center column). This synchronization change occurs as the power spectrum of the stimulus is shifted such that the effective BF is moved from the second formant of /I/ toward the second formant of /æ/.

B. Primarylike units

The responses of a pri-N unit are shown in Fig. 4, which is organized in the same fashion as Fig. 3. The top row of Fig. 4 shows the response to the steady-state vowel portion of the first stimulus /i/ (98 Hz). Pri-N units differ from choppers in that they phase lock well to frequencies up through the second and third formant frequency range (Blackburn and Sachs, 1990); there is a cluster of frequency components near the vowel's second and third formants in the response shown in the right column. However, there is also strong synchronization to F_0 at 98 Hz (center column) and its harmonics. The bottom row shows the response to the steady-state vowel portion of the second syllable, /ʌ2/ (112 Hz). Again strong synchronization can be seen at both F_0 and its harmonics (center column) and centered around the third formant (right column). The second through sixth rows show responses to the vowel pairs as F_P is increased from 50 to 64 kHz. This change in F_P moves the unit's effective BF from near the third formant of /i/ (98 Hz) to near the third formant of /ʌ2/ (112 Hz). Notice that the F_0 synchronization changes from predominantly locked to base-rate 98 Hz to predominantly locked to base-rate 112 Hz over this range of F_P.

The data in Figs. 3 and 4 are typical of the responses of our population of chopper and primarylike units. Similar results were obtained for units over the range of BFs studied (0.69–5.27 kHz). These data show clearly the basic behavior hypothesized for the periodicity-tagged spectral representation, which is that the temporal response pattern reflects the fundamental frequency of the stimulus whose components dominate near the BF of the unit. Surprisingly, similar behavior was observed in both chopper and primarylike units. In the next section, we examine the second part of the hypotheses, that the rate response of a unit provide a monotonic measure of the energy in the stimulus within the unit's tuning curve.

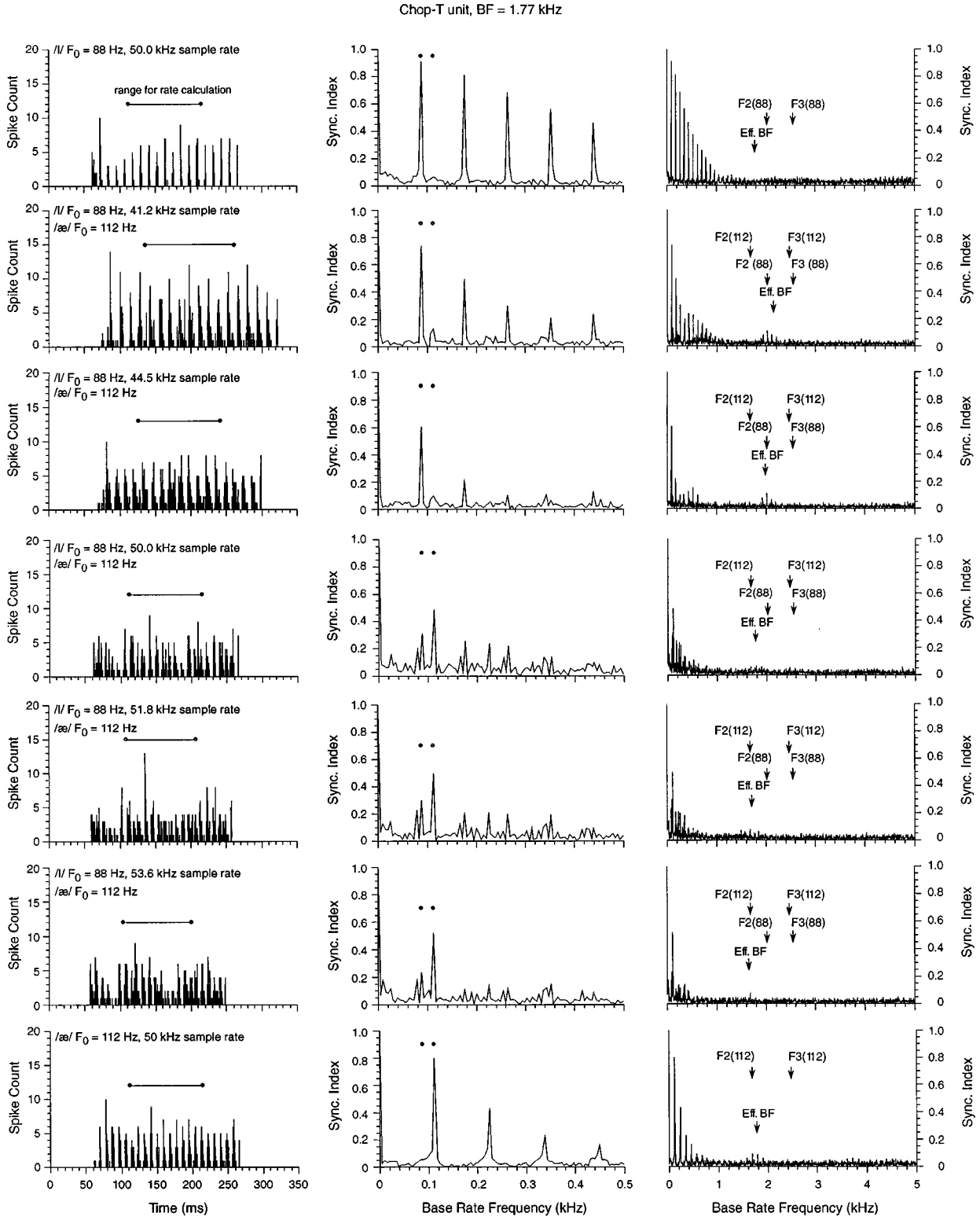


FIG. 3. Responses of a chop-T unit (BF=1.77 kHz) to single- and double-vowel stimuli; the vowels are identified in the legends in the left column, along with F_P. The top and bottom rows show responses to the two vowels in isolation and the second through sixth rows show responses to double-vowel stimuli with various values of F_P. The first column plots the PSTH for the steady-state vowel portion only of the syllables (from 61.44 to 266.24 ms when the playback sampling rate is equal to the synthesis sampling rate, Kaiser window not yet applied), and the second and third columns plot the discrete Fourier transform of the portion of the PSTH associated with the steady-state vowel (Kaiser window applied). The abscissae of the Fourier transforms are plotted in terms of base-rate frequency, explained in the text, b_k in Eq. (4). Frequencies near the fundamental are shown in the center column and the full frequency range is plotted in the right column. The horizontal bars in the left column show the portion of the response from which average rate was computed for the analysis in Fig. 6. The dots in the center column mark the two F_0's (88 and 112 Hz) and the formant frequencies of the vowels are shown by the arrows in the right column. Note that the F_0's and formant frequencies do not change as the sampling rate changes because the Fourier transforms are plotted in terms of base-rate frequency. The BFs (top and bottom plot where base-rate frequency equals true frequency) and effective BFs are also shown in the right column.

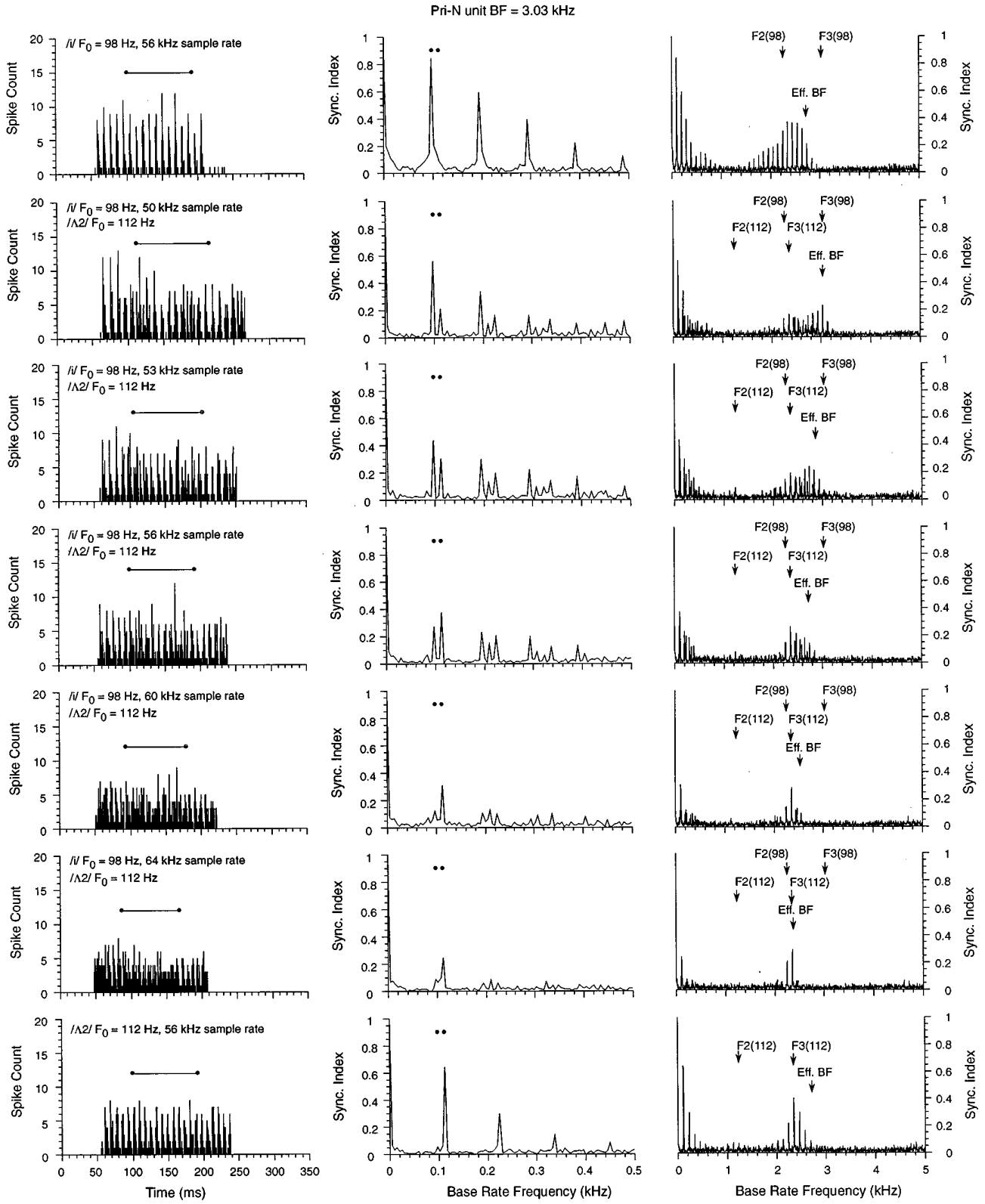


FIG. 4. Responses of a pri-N unit (BF=3.03 kHz) to single- and double-vowel stimuli. Figure layout is the same as for Fig. 3.

III. RESULTS: RATE RESPONSES

A. Analysis methods

In this section, we analyze the average discharge rates of units for sequences of stimuli like those shown in Figs. 3 and 4. Rate will be compared with the stimulus energy near a unit's BF. To provide an estimate of the stimulus energy ''seen'' by a unit, the double vowel's power spectrum was filtered by the unit's frequency tuning curve (TC). This pro-

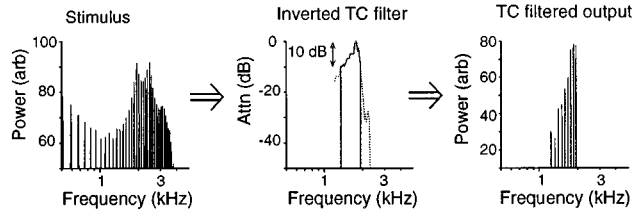


FIG. 5. The filtering scheme used to derive measures of the stimulus energy near a unit's BF. The stimulus power spectrum is shown at left and the unit's inverted tuning curve is shown in the middle. The solid-line portion of the tuning curve shows gains within 10 dB of the peak; the dashed portion of the tuning curve, where gains are smaller, was set to 0 gain and ignored during the filtering. The filtered stimulus is shown at right. Note that the ordinate scale is arbitrary. Two measures of stimulus energy, P_{10} and $P_{\max}$, were derived from the filtered output, as described in the text.

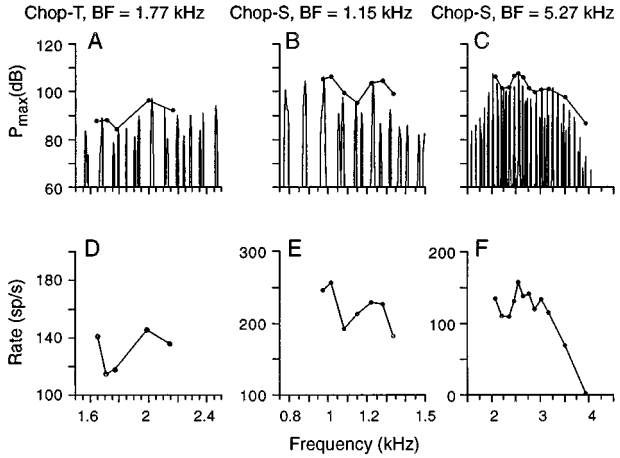


FIG. 6. Comparison of $P_{\max}$ (top row) and average discharge rate (bottom row) for three chopper units. (A) Spectrum of the double vowel stimulus /bɪʃ/ ($F_0=88$ Hz) and /æʃ/ ($F_0=112$ Hz) plotted against base-rate frequency. The superimposed line with filled circles shows $P_{\max}$ computed from this double-vowel stimulus and the TC of a chop-T unit (same unit as Fig. 3) for 5 values of F_P. $P_{\max}$ is plotted against effective BF. (B) Same as (A) for a chop-S unit and the double-vowel /buʃ/ ($F_0=100$ Hz) and /bΛ2ʃ/ ($F_0=88$ Hz). (C) Same as (A) for another chop-S unit and the double-vowel /bɪʃ/ ($F_0=88$ Hz) and /biʃ/ ($F_0=112$ Hz). (D)–(F) Average discharge rates of the three units in response to the double vowel stimuli, plotted against effective BF. The $P_{\max}$ plots in (A), (B), and (C) were calculated for the same F_P's as the average rate plots in (D), (E), and (F).

cess is illustrated in Fig. 5. The filter used was the inverted tuning curve of the unit, with its peak value set to 0-dB gain. Two measures of the output of the filter are described below. These measures are arbitrarily chosen, but emphasize energy near the BF of the unit because it is generally found that, in the AN, the responses of units are either dominated by stimulus components near their BFs or by the stimulus component at the nearest formant peak (Young and Sachs, 1979; Sinex and Geisler, 1983; Delgutte and Kiang, 1984a, b; Palmer *et al.*, 1986; Palmer, 1990). The measures used are as follows:

(1) The total power in the filtered stimulus over the frequency range where the TC gain is within 10 dB of its maximum (solid line in Fig. 5), called P_{10} below.
(2) The power in the single largest component within the 10-dB bandwidth (after filtering), called $P_{\max}$ below.

This procedure was applied to both single-vowel and double-vowel stimuli and was repeated for each sampling rate, yielding a different P_{10} and $P_{\max}$ for each effective BF.

B. Rate responses of chopper units

Figure 6 shows a comparison of average discharge rate and stimulus power near a unit's BF as measured by $P_{\max}$. Each column shows data from a different chopper unit; the data in the left column are from the same unit as in Fig. 3. Figure 6A–C shows the spectrum of a double-vowel stimulus, plotted against base-rate frequency, with a superimposed plot of $P_{\max}$ versus effective BF for the F_P's used for that unit. As expected, there is a general correspondence between the envelope of the stimulus spectrum and $P_{\max}$. A similar correspondence is seen when P_{10} is plotted (not shown). Figure 6D–F shows the unit's average discharge rates in response to the same stimuli over the same range of F_P's, plotted against effective BF. Notice the qualitative similarities between the rate measurements (bottom row) and the $P_{\max}$ values (top row). In order to quantify this similarity, correlation coefficients were calculated between the average rate data and both $P_{\max}$ and P_{10} plots. The correlation coefficients were translated to Fisher-z scores, and an average z-score, weighted by the different number of measurements for each unit, was computed across all data sets for each unit type. Finally, the average z-score was then converted back to a correlation coefficient. The resulting average correlation coefficients for chop-S and chop-T units are 0.86 and 0.62 respectively, for $P_{\max}$ and 0.88 and 0.78 for P_{10}. All values are significantly different from zero (with a z-test for z-score, $p<0.05$ for all cases). This analysis verifies the second part of our hypothesis, by showing that chopper units encode the average energy in the stimulus near their BFs.

C. Synchronization of chopper units to F_0

Figure 7 shows an analysis of chopper units' synchronization to the F_0's of the vowel stimuli. Data are shown for the same three units as in Fig. 6. Figure 7D–F shows the synchronization indices of the units' responses at the three F_0's, plotted against effective BF. In each case, the F_P's carry the effective BF of the unit from a point near one vowel's formant to a point near the other vowel's formant, as in Figs. 3 and 4. Note that one of the three F_0's in each case was not physically present in the signal and shows near-zero synchronization; the third F_0 is plotted as an indication of the noise level of the analysis. These data show the same trend as in Figs. 3 and 4: There is a transition from dominance of the units' temporal responses by one vowel to dominance by the other as the effective BF moves from one formant to the other. We define the *synchronization crossing* point as the point at which the synchronizations to the two F_0's are equal (arrows).

A reasonable prediction of the synchronization crossing points is provided by the tuning-curve analysis shown in Fig. 7A–C. For this analysis, the two vowels making up the double-vowel stimulus were filtered separately (i.e., as single

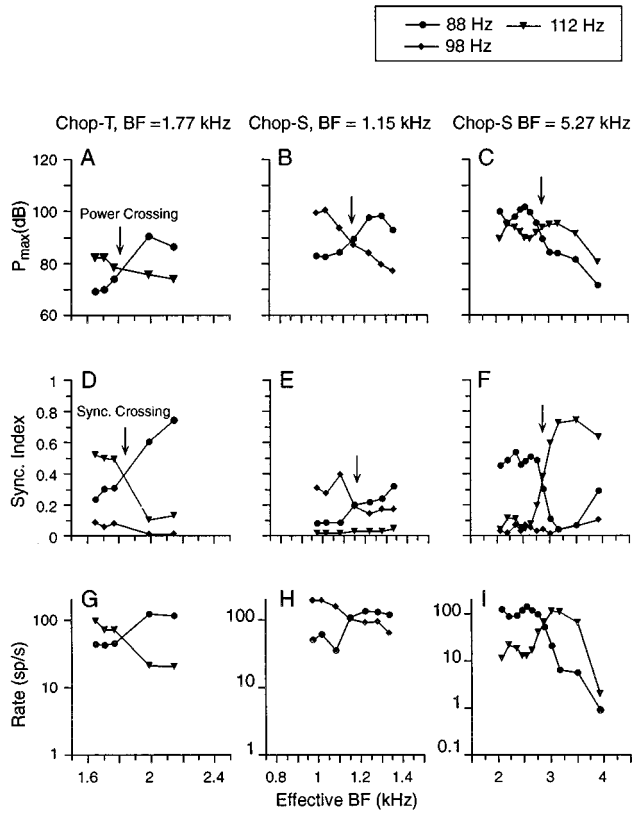


FIG. 7. Analysis of synchronization for three chopper units (same units and same double-vowel stimuli as in Fig. 6). Each column shows data from one unit. (A), (B), and (C) $P_{\max}$, computed separately from the single vowels making up the double-vowel stimulus, plotted versus effective BF. Symbols correspond to different vowels, identifiable by their F_0's, given in the legend; vowels are listed in the caption of Fig. 6. Arrows point to power crossings, where the $P_{\max}$ for each unit's TC has the same value for both components of the double vowel. (D), (E), and (F) Synchronization indices at the three possible F_0's, computed from the responses of these units to double vowels over a range of F_P's. In each case, only two of the three F_0's are actually present in the stimulus, so that the power at the third frequency is noise. (G), (H), and (I) Segregated rates for the two vowels making up the double-vowel, computed using Eq. (5) from the average rates in Fig. 6D–F. These are plotted on a logarithmic ordinate to make the scaling comparable to the dB scale for $P_{\max}$ in A–C.

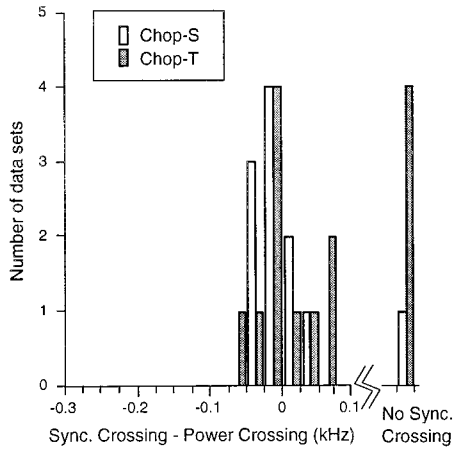


FIG. 8. Histogram showing the distribution of the difference between the effective BFs of the synchronization crossings (e.g., Fig. 7D) and power crossings (e.g., Fig. 7A). Data are shown for all chop-S (5 units, 14 data sets) and chop-T (8 units, 11 data sets) units that had a power crossing. Bins beyond the axis break show data sets with a power crossing but no synchronization crossing. Four data with neither power nor synchronization crossings are not shown.

vowels) through the unit's tuning curve to compute $P_{\max}$ and P_{10} for the playback sampling rates used for the data in Fig. 7D–F. Similar results were obtained for $P_{\max}$ and P_{10} and results are shown in Fig. 7A–C for $P_{\max}$ only. Note the qualitative similarity of the $P_{\max}$ calculations (Fig. 7A–C) and the synchronization data (Fig. 7D–F). The *power crossing* is defined as the point at which the $P_{\max}$'s of the two vowels are equal (arrows). On either side of the power crossing one of the vowels has greater power than the other and should dominate both the double vowel spectrum and the unit's response at that effective BF. In agreement with this expectation, there is a good correspondence between the power and synchronization crossing frequencies.

The histogram in Fig. 8 shows the distribution of the difference in frequency between synchronization and power crossings for all the chopper data sets. Similar results were obtained for the $P_{\max}$ and P_{10} measures. The differences for both chop-S and chop-T units are centered around zero, showing that the point at which the synchronization indices cross was reasonably predicted by the power crossings. The average difference was 0.0 Hz (s.d. 40.0 Hz) for chop-S units (10 data sets) and -8.9 Hz (s.d. 24.7 Hz) for chop-T units (10 data sets). Four data sets from one chop-S unit and 1 data set from a chop-T unit had power crossings but no synchronization crossings. These are not included in the average differences given above but are shown on the histogram after the axis break. Two chop-S and two chop-T data sets had neither power nor synchronization crossings. The results shown in Figs. 7 and 8 quantify the qualitative conclusions drawn from Fig. 3.

D. Separation of double vowels in chopper units

In order to evaluate whether VCN units provide the necessary information to allow segregation of multiple sound sources, we investigated the hypothesis that the relative synchronization to the two F_0's of a double vowel can be used to apportion the unit's discharge rate between the two vowels, and therefore segregate them. We define an estimate R_{V1} of the portion of the total discharge rate R_{total} of a unit that should be assigned to one vowel of a double-vowel pair:

$$R_{V1} = R_{\text{total}} \frac{S(F_{0\,V1})}{S(F_{0\,V1}) + S(F_{0\,V2})}, \tag{5}$$

where $S(F_{0\,V1})$ and $S(F_{0\,V2})$ are the synchronization indices of the unit's response to the two vowels' F_0's. The portion of the rate assigned to vowel 2 (R_{V2}) is computed similarly. Note that the same results would be obtained if synchronized rates, i.e., $P(F_{0\,Vi})$ in Eq. (3), were substituted for synchronization $S(F_{0\,Vi})$ in Eq. (5).

If the segregated rates R_{V1} and R_{V2} are similar to the unit's discharge rates in response to the vowels $V1$ and $V2$ presented singly, then we have shown that there exists an algorithm which can reconstruct the individual constituents

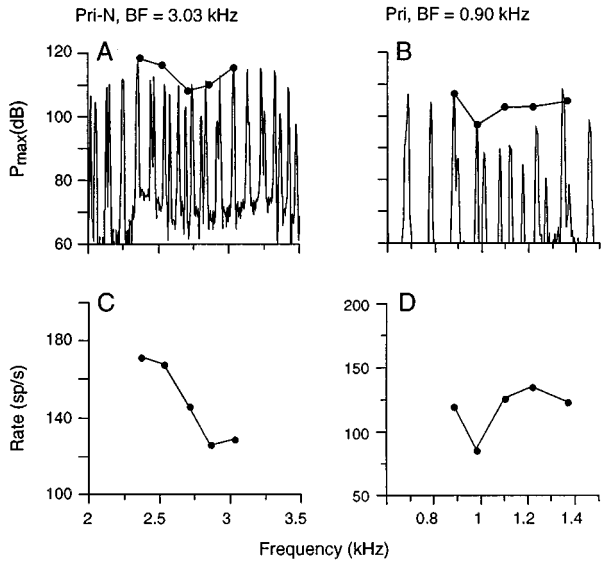


FIG. 9. Same analyses as in Fig. 6 for two primarylike units. (A) $P_{\max}$ from the output of the TC filtering is shown overlaid on the spectrum of the double-vowel /biʃ/ ($F0=98$ Hz) and /bʌ2ʃ/ ($F_0=112$ Hz) for a pri-N unit, BF=3.03 kHz (same unit as in Fig. 4). (B) Same for a pri unit, BF=0.9 kHz, with the double-vowel /bɔʃ/ ($F0=98$ Hz) and /bɜʃ/ ($F0=112$ Hz). $P_{\max}$ is plotted vs effective BF and the spectra are plotted vs base-rate frequency. (C), (D) Average discharge rate as a function of effective BF for the same units and stimuli.

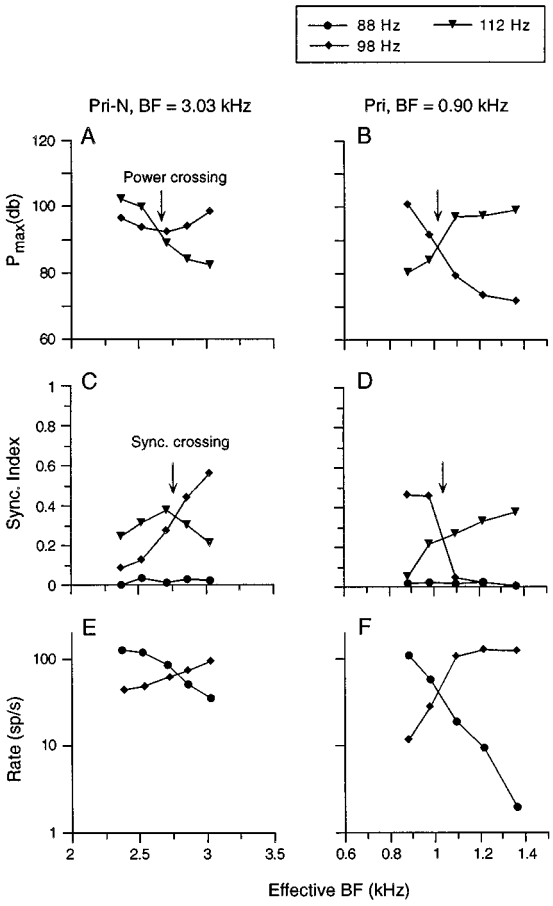


FIG. 10. Same analyses as in Fig. 7 for primarylike units; same units as in Fig. 9.

of the double-vowel stimulus. Unfortunately, we did not take data on the responses to single vowels at F_P's other than the synthesis rate; thus the R_{Vi} must be compared with an indirect measure. A convenient measure is $P_{\max}$ as plotted in Fig. 7A–C. Note that the R_{Vi} and $P_{\max}$ have different units, so only a qualitative comparison is possible.

Figure 7 G–I shows plots of the segregated rates R_{V1} and R_{V2} for the three example units. These figures should be compared with Fig. 7A–C which shows the corresponding $P_{\max}$ values. The R_{Vi} are plotted on a logarithmic ordinate in Fig. 7G–I to provide similar scaling as for the $P_{\max}$ data in Fig. 7A–C. The important observations from this comparison are as follows: First, the crossings observed in the $P_{\max}$ plots are preserved in the segregated rate data. In fact the frequencies of the crossing points in the segregated rate data are virtually identical to the frequencies of the synchronization crossings, because of the way the segregated rates are computed [Eq. (5)]. Thus the correspondence of segregated-rate crossings and power crossings follows from the correspondence of the power and synchronization crossings. Second, the general shape of the segregated rate plots matches that of the TC-filtered power plots. Third, the formant peaks of the individual vowels can, in most cases, be mapped to the peaks of the segregated rate measure. This correspondence can be seen by comparing the effective BFs of the peaks in Fig. 7A–C, which corresponds almost exactly with the formant frequencies, to the peaks in the segregated rates in Fig. 7G–I. The correspondence of peaks is not as close when synchronization index plots (Fig. 7D–F) are used.

E. Primarylike units

Figures 9–11 show the same analyses as Figs. 6–8 for pri and pri-N units. Figure 9 compares $P_{\max}$ with average discharge rate for two primarylike units. The filtered stimu-

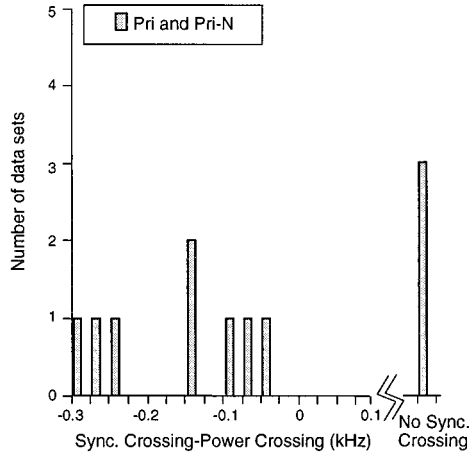


FIG. 11. Histogram of the difference between synchronization and power crossings for all primarylike units (11 data sets from 7 units). Bars beyond the break in the abscissa show cases with power crossing but no synchronization crossing. Five data sets with neither power nor synchronization crossings are not shown.

lus and the rate measure for the unit in Fig. 9B and 9D are similar, in that they share a dip at 1 kHz, however, the curves are rather different for the unit in Fig. 9A and 9C. The correlation analysis described above for choppers was applied to the primarylike data and yielded average correlation coefficients between the filtered stimulus and the rate measure of 0.35 for $P_{\max}$ and 0.41 for P_{10} (pri and pri-N were combined for this analysis). While these scores are lower than those found for the chopper units they are still statistically distinct from zero ($p<0.05$, z-test for z-score for all cases). Thus, the pri and pri-N units are also able to encode information regarding the overall double-vowel stimulus envelope.

Figure 10A–B shows the $P_{\max}$ values for the two primarylike units for the two vowels separately and Fig. 10C–D shows the synchronization to the fundamental frequencies of the vowels. Power and synchronization crossings are seen in both units and appear to correspond well in these two examples. Figure 11 shows a histogram of the differences in effective BF between the synchronization and power crossings for all primarylike units; the average difference is -156 Hz (s.d. 99 Hz; 8 data sets from both pri and pri-N units) showing a bias towards lower frequency in the synchronization crossing compared to the power crossing. Not included in the average are three data sets which had power crossings but no synchronization crossings; these are shown in the histogram beyond the axis break. Five additional primarylike unit data sets had neither a synchronization nor a power crossing. Thus the F_0 synchronization of primarylike units contains information regarding the power crossings but that information is not as accurate as in the chopper units.

The segregated rates for these units and stimuli are shown in Fig. 10E–F. The general appearance of these two examples is similar to that of the choppers in Fig. 7. However, from the synchronization data summarized above, and from the similarity in crossing points in synchronization and segregated-rate data, we expect that primarylike units will not provide as good information for source segregation as chopper units.

IV. DISCUSSION

A. Periodicity-tagged spectral representation in choppers and primarylike units

The results reported above are consistent with the idea that VCN chopper units, and to a lesser extent primarylike units, provide a representation of multiple-source speech stimuli in which the stimulus spectrum is encoded in discharge rate and the fundamental frequency of the source is encoded in the temporal patterns of discharge of the units. The representation is clearest in the case of chopper units with BFs above 500–700 Hz, where the temporal response patterns reflect only phase locking at the F_0's. That is, these choppers do not provide a temporal representation of second and third formants, but they do provide a rate representation of the stimulus spectrum and a temporal representation of F_0, as required for the periodicity-tagged representation. Our results extend Blackburn and Sachs' (1990) results for single-vowel stimuli to the case of double-vowel stimuli; consistent with their findings, we have shown that choppers provide a good rate representation of these more complex spectra, at least over the narrow range of stimulus levels examined here.

At BFs below 500–700 Hz, the idea of a periodicity-tagged spectral representation in choppers is more uncertain. We did not successfully record from choppers at these low BFs, but some inferences can be made about the responses of low BF units from existing data. Based on the single-vowel data (Blackburn and Sachs, 1990), these neurons should give a good rate representation of the stimulus spectrum and their temporal responses should be phase locked to both F_0 and to harmonics near the first formant. Thus, the simple periodicity-tagged representation seen in high BF choppers should be complicated, in low BF choppers, by phase locking to stimulus components near the first formants of a double vowel. In fact, the responses of low BF choppers should resemble those of primarylike neurons with BFs near the second and third formants (discussed below), by showing a mixed phase locking to F_0 and to harmonics near the formants. However, even with the more complex phase locking, there should be a substantial component of response at F_0 (e.g., Fig. 10A and 10B of Blackburn and Sachs, 1990) and it should be possible to derive the same kind of information from low BF choppers as from high BF choppers, albeit with a more complex algorithm. However, the important question of what information neural periodicity can provide about vowel identity in the lower half of the first-formant frequency region remains to be determined.

Primarylike units provide both a rate and a temporal representation of the formant frequencies of a single vowel (Palmer *et al.*, 1986; Blackburn and Sachs, 1990; Winter and Palmer, 1990b). Although it has not been demonstrated experimentally, it seems plausible from the single-vowel data that primarylike units provide as good a temporal representation of double-vowel stimuli as is observed in the AN (Palmer, 1990). Thus the nature of the representation of double vowels in the population of primarylike units is more complex than in the choppers. At minimum, the primarylike temporal representation contains information about both F_0 and the formants. The complexity of primarylike phase locking may explain in part why the quality of the periodicity-tagged spectral representation is weaker in primarylike units, compared to choppers (e.g., Fig. 11). In addition, primarylike units generally provide a weaker rate representation of stimulus spectrum than do choppers (Blackburn and Sachs, 1990).

Because primarylike units provide a robust temporal representation of vowel spectra, the quality of the primarylike periodicity-tagged representation is expected to be better if evaluated in temporal rather than in rate measures. Although temporal representations of the double-vowel spectra were not evaluated in this paper, examination of Fig. 4 shows clearly the behavior expected of such a temporal representation; from the second to the sixth rows in this figure, the units' phase locking goes from being centered on the third formant of the /i/ (98 Hz), when effective BF is near the /i/'s third formant, to being centered on the third formant of the /ʌ2/ (112 Hz), when effective BF is near the /ʌ2/'s third formant. Use of the temporal information in primarylike units would, of course, require the brain to segregate phase

locking to F_0 from phase locking to the formants.

A related issue, in the case of primarylike units, is that synchronization to frequency F_0 is an incomplete measure of the degree of response modulation at frequency F_0, for units that phase lock strongly at higher frequencies. In such units, envelope modulation at F_0 can result from a component at F_0 itself or from any group of unresolved harmonics of F_0, such as the ones centered on the unit's BF (e.g., Fig. 4, right column; Wang and Sachs, 1993; Miller and Sachs, 1984). Indeed, it is theoretically possible to have substantial modulation of a neuron's discharge rate at frequency F_0 with essentially no synchronization at frequency F_0. Thus it is possible that we obtained poorer results in primarylike than in chopper units because we used an inappropriate measure for estimating the degree of response modulation by F_0 in primarylike units. That this is not the case is suggested by two facts: (1) Wang and Sachs (1993) showed, for AN fibers, that an estimator of envelope modulation which includes all sources of modulation usually varies monotonically with synchronization at F_0. Thus even though Wang and Sachs' measure of envelope modulation has properties which make it preferable to F_0 synchronization, it is unlikely to give different answers for the analyses done here, as long as significant synchronization to F_0 is observed. (2) In the primarylike units in our data sample, significant synchronization to F_0 was observed (Figs. 4 and 7), which is somewhat different from the result in the AN. Presumably, the F_0 synchronization in primarylike units results from nonlinearities in synaptic processing which generate F_0 components as difference tones of unresolved harmonics in the AN input. In any case, it was necessary to use F_0 synchronization as a measure of response modulation at F_0 in this paper, because with a double-vowel stimulus with more than one F_0 modulation, it is not possible to use methods like that of Wang and Sachs (1993).

B. Periodicity-tagged spectral representation as a basis for source segregation

The representation described here seems to contain sufficient information to support segregation of the components of a double vowel on the basis of differences in F_0. F_0 as a cue for segregating the energy in a double vowel is strongly suggested by the body of perceptual evidence (e.g., Assmann and Summerfield, 1990, 1994; Culling and Darwin, 1993, 1994). We have provided a simple scheme for carrying out the segregation [Eq. (5)]. Although this scheme is an artificial construct, it does demonstrate the possibilities inherent in this representation.

The representation we have demonstrated is similar in concept to the first stage of many of the models of double-vowel segregation (Parsons, 1976; Assmann and Summerfield, 1990; Meddis and Hewitt, 1992), in which an excitation pattern is computed and the F_0's of the vowels present in the stimulus are estimated. In the second stage of the models, the F_0's are used to guide samplings of the excitation pattern in order to segregate the energy in the two vowels, so that a template-matching process can identify them. In our case, the F_0 is represented in an encoded form and it is interesting to speculate on how the F_0 information could be used in a neural second stage. There are two possibilities: (1) The F_0 could be detected by neurons which only respond to certain temporal patterns in their inputs. If a range of temporal sensitivities were available, then the two vowels of the stimulus would end up in different places along an F_0 map. At that point, the F_0 would have been converted from a temporal to a place representation. Evidence for the existence of such a map of fundamental frequency has been provided in the inferior colliculus where units have sensitivities to F_0 that vary from 10 to 1000 Hz. (Schreiner and Langner, 1988). (2) The F_0 could be maintained as the marker of the response so that the energy in one of the two vowels would be distributed across an array of central neurons, tagged by its F_0. In this model, the components making up a vowel would be associated together centrally by their common temporal patterns. Although this sort of distributed representation is attractive from the standpoint of not requiring special mechanisms and maps at higher levels, the evidence argues against it, in that neurons in the auditory cortex do not encode periodicity in their temporal response patterns at frequencies above 10–20 Hz (reviewed by Langner, 1992). Thus the second representation would still have to be converted to a place representation at a subcortical level.

C. Effects of stimulus level

Because the time required to obtain one of the data sets used here approaches the limited time over which contact with a unit can be maintained, we did not explore a broad range of stimulus levels. The stimulus levels chosen in these experiments are, in a sense, optimal, in that they were chosen to avoid saturation. However, in the case of chopper units, Blackburn and Sachs (1990) showed that saturation of the rate-level function for BF tones does not imply saturation of the units' responses to broadband stimuli. Presumably choppers avoid saturation by some sort of inhibitory process; this could include lateral inhibition for enhancing stimulus contrast, but there are other possibilities. Lai and colleagues (1994a, b) have developed a model of CN choppers that shows saturating rate functions for BF tones but still allows the units to have a wide dynamic range for stimuli such as speech; the model is based on the idea (Winslow *et al.*, 1987) that shunting inhibition in the dendritic tree can be used to switch a stellate neuron's inputs from high-spontaneous rate AN fibers to low-spontaneous rate fibers. The switch is arranged so that high-spontaneous rate fibers dominate at low sound levels and low-spontaneous rate fibers take over at high levels. Such a switch of inputs provides a simple explanation of the robustness of the rate responses of CN chopper units to single-vowel stimuli (Blackburn and Sachs, 1990); evidence for such a mechanism is provided by the fact that chopper rate responses to speech superimpose nicely on high-spontaneous rate AN fibers' rate responses at low sound levels and on low-spontaneous rate fibers' responses at high sound levels. In our case, we found no evidence for a degeneration of the periodicity-tagged representation with sound level, in that there was no difference in the representation at 30 versus 20 dB *re*: threshold, but conclusions about the robustness of the

periodicity-tagged spectral representation must await data on the effects of stimulus level over a wider range.

The stimulus levels we used are optimal in the additional sense that the responses of a unit to the two components of a double-syllable were approximately equal, when the components were presented separately at the base playback rate. This matching made it likely that each vowel's harmonics dominated the double vowel's spectrum over some frequency range near its formants. As one of the two syllables is made more intense than the other, a point is reached where the harmonics of the weaker syllable never dominate within the unit's tuning curve. In this case, the responses of the unit should never be dominated by the weaker vowel. The weak-vowel/strong-vowel situation occurred in our data in those cases where neither power nor synchrony crossings were observed (four chopper units and five primarylike units; see the discussion of Figs. 8 and 11). When one vowel is sufficiently weak compared to the other, chopper units along the whole tonotopic axis would have temporal responses (and rate responses) which reflect mainly the stronger vowel. Any representation of the weaker vowel that remained might be extracted by the rate segregation scheme of Eq. (5), but it is is not clear how good a representation of the spectrum of the weaker vowel could be derived. A good test of the periodicity-tagged representation would be to compare its performance as the relative level of the two components of a syllable varied with the perceptual masking of the weaker syllable by the stronger.

D. Phase locking to F_0 in cochlear nucleus

The periodicity-tagged spectral representation we propose depends on a CN unit phase locking to the fundamental frequency of the vowel whose formant dominates in the tuning curve of the unit; what this means is that the unit should phase lock to the F_0 of a vowel when the unit's BF is near one of the vowel's formant peaks. The source of this phase locking requires some comment, based on the properties of AN fibers. AN fibers whose BFs are near formant peaks tend to phase lock strongly to the formant component (for synthetic vowels with formants corresponding exactly to a harmonic, as used here). As stimulus level increases, phase locking to the formant dominates the response, and phase locking to other frequency components, including F_0, is suppressed (Young and Sachs, 1979; Miller and Sachs, 1984; Delgutte and Kiang, 1984a; Geisler and Silkes, 1991; Wang and Sachs, 1993); accompanying this suppression is a loss of modulation of the unit's response at frequency F_0. The same suppression of synchronization to F_0 is observed in responses to double vowels (Palmer, 1990). These data suggest that the rate modulation at frequency F_0 in the input to a CN neuron should be weak, when the CN neuron's BF is near a formant. Thus some amplification of the F_0 response appears to be necessary in CN units.

In fact, a number of studies have shown that there is an amplification of the representation of the temporal envelope of the stimulus in going from AN to CN neurons (Frisina *et al.*, 1990a; Kim *et al.*, 1990; Wang and Sachs, 1994; Rhode and Greenberg, 1994). Onset neurons generally show the largest effect, followed by choppers and pri-N units, with pri units having essentially no amplification. Several mechanisms for the amplification have been discussed in both single-unit and modeling studies (Frisina *et al.*, 1990b; Wang and Sachs, 1994, 1995). Wang and Sachs (1995) analyzed a model of chopper units; the phase locking properties of this model are determined by dendritic low-pass filtering, which prevents the unit from phase locking to stimulus components above a few hundred Hz and thereby isolates the F_0 component in the response. Wang and Sachs showed that amplification of the F_0 component could occur by convergence of subthreshold inputs, by convergence of inputs over a range of BFs, and by adjustment of the threshold of the model unit's dynamic range with inhibition. Similar studies have not been done on pri-N or onset units.

Regardless of mechanism, the existence of enhanced temporal envelope sensitivity in the cochlear nucleus suggests a role for time-domain analysis in acoustic signal processing. Our data suggest that one aspect of this role could be as an initial stage of processing for F_0-based segregation of simultaneous speech streams.

V. SUMMARY AND CONCLUSIONS

When two steady-state vowels are concurrently presented, chopper units of the VCN alter their average discharge rates in accordance with the spectral power of the total stimulus within the units' tuning curves. A chopper's responses are synchronized to either or both of the vowels' fundamental frequencies, depending on the relative strengths of the vowels in the unit's tuning curve. The relative strength of synchronization can be used to segregate the rate response of a unit between two vowels in a demonstration that chopper units carry the information necessary for the segregation of competing auditory streams. Similar results were obtained for primarylike units, although the effects were less pronounced than for chopper units.

ACKNOWLEDGMENTS

The authors would like to thank P. P. Taylor for her assistance with the figures. The journal's reviewers, Alan Palmer and Chris Darwin, made suggestions that substantially improved the manuscript. This research was supported by the Johns Hopkins Center for Language and Speech Processing and by grants from NIDCD (DC00109 and DC00023).

Assmann, P. F., and Summerfield, Q. (**1989**). ''Modeling the perception of concurrent vowels: Vowels with the same fundamental frequency,'' J. Acoust. Soc. Am. **85,** 327–338.

Assmann, P. F., and Summerfield, Q. (**1990**). ''Modeling the perception of concurrent vowels: Vowels with different fundamental frequencies,'' J. Acoust. Soc. Am. **88,** 680–697.

Assmann, P. F., and Summerfield, Q. (**1994**). ''The contribution of waveform interactions to the perception of concurrent vowels,'' J. Acoust. Soc. Am. **95,** 471–484.

Blackburn, C., and Sachs, M. B. (**1989**). ''Classification of unit types in the anteroventral cochlear nucleus: PST histograms and regularity analysis,'' J. Neurophysiol. **62,** 1303–1329.

Blackburn, C., and Sachs, M. B. (**1990**). ''The representations of the steady-state vowel sound /e/ in the discharge patterns of cat anteroventral cochlear nucleus neurons,'' J. Neurophysiol. **63,** 1191–1212.

Bourk, T. R. (**1976**). ''Electrical responses of neural units in the anteroventral cochlear nucleus of the cat,'' Ph.D. thesis, MIT, Cambridge.
Brokx, J. P. L., and Nooteboom, S. G. (**1982**). ''Intonation and the perceptual separation of simultaneous voices,'' J. Phon. **10**, 23–36.
Bronkhurst, 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.
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-Verlag, New York), pp. 66–116.
Conley, R. A., and Keilson, S. E. (**1995**). ''Rate representation and discriminability of second formant frequencies for /ε/-like steady-state vowels in cat auditory nerve,'' J. Acoust. Soc. Am. **98,** 3223–3234.
Culling, J. F., and Darwin, C. J. (**1993**). ''Perceptual separation of simultaneous vowels: Within and across-formant grouping by F_0,'' J. Acoust. Soc. Am. **93,** 3454–3467.
Culling, J. F., and Darwin, C. J. (**1994**). ''Perceptual and computational separation of simultaneous vowels: Cues arising from low-frequency beating,'' J. Acoust. Soc. Am. **95,** 1559–1569.
Darwin, C. J. (**1981**). ''Perceptual grouping of speech components differing in fundamental frequency and onset time,'' Q. J. Exp. Psychol. **33A**, 185–207.
Darwin, C. J. (**1984**). ''Perceiving vowels in the presence of another sound: Constraints on formant perception,'' J. Acoust. Soc. Am. **76,** 1636–1647.
de Cheveigné, A. (**1993**). ''Separation of concurrent harmonic sounds: Fundamental frequency estimation and a time-domain cancellation model of auditory processing,'' J. Acoust. Soc. Am. **93,** 3271–3290.
de Cheveigné, A., McAdams, S., Laroche, J., and Rosenberg, M. (**1995**). ''Identification of concurrent harmonic and inharmonic vowels: A test of the theory of harmonic cancellation and enhancement,'' J. Acoust. Soc. Am. **97,** 3736–3748.
Delgutte, B., and Kiang, N. (**1984a**). ''Speech coding in the auditory nerve: I. Vowel-like sounds,'' J. Acoust. Soc. Am. **75,** 866–878.
Delgutte, B., and Kiang, N. (**1984b**). ''Speech coding in the auditory nerve: II. Processing schemes for vowel-like sounds,'' J. Acoust. Soc. Am. **75,** 878–886.
Delgutte, B., and Kiang, N. (**1984c**). ''Speech coding in the auditory nerve: III. Voiceless fricative consonants,'' J. Acoust. Soc. Am. **75,** 887–896.
Delgutte, B., and Kiang, N. (**1984d**). ''Speech coding in the auditory nerve: I. Sounds with consonantlike dynamic characteristics,'' J. Acoust. Soc. Am. **75,** 897–907.
Frisina, R. D., Smith, R. L., and Chamberlain, S. C. (**1990a**). ''Encoding of amplitude modulation in the gerbil cochlear nucleus. I. A hierarchy of enhancement,'' Hearing Res. **44,** 99–122.
Frisina, R. D., Smith, R. L., and Chamberlain, S. C. (**1990b**). ''Encoding of amplitude modulation in the gerbil cochlear nucleus. II. Possible neural mechanisms,'' Hearing Res. **44,** 123–142.
Geisler, C. D., and Silkes, S. M. (**1991**). ''Responses of 'lower-spontaneous-rate' auditory-nerve fibers to speech syllables presented in noise. II: Glottal-pulse periodicities,'' J. Acoust. Soc. Am. **90,** 3140–3148.
Johnson, D. H. (**1980**). ''The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones,'' J. Acoust. Soc. Am. **68,** 1115–1122.
Keithley, E. M., and Schreiber, R. C. (**1987**). ''Frequency map of the spiral ganglion in the cat,'' J. Acoust. Soc. Am. **81,** 1036–1042.
Kim, D. O., Sirianni, J. G., and Chang, S. O. (**1990**). ''Responses of DCN-PVCN neurons and auditory nerve fibers in unanesthetized decerebrate cats to AM and pure tones: Analysis with autocorrelation/power-spectrum,'' Hearing Res. **45,** 95–114.
Klatt, D. H. (**1980**). ''Software for a cascade/parallel formant synthesizer,'' J. Acoust. Soc. Am. **67,** 971–995.
Lai, Y.-C., Winslow, R. L., and Sachs, M. B. (**1994a**). ''The functional role of excitatory and inhibitory interactions in chopper cells of the anteroventral cochlear nucleus,'' Neural Comput. **6,** 1127–1140.
Lai, Y.-C., Winslow, R. L., and Sachs, M. B. (**1994b**). ''A model of selective processing of auditory-nerve inputs by stellate cells of the anteroventral cochlear nucleus,'' J. Comput. Neurosci. **1**, 167–194
Langner, G. (**1992**). ''Periodicity coding in the auditory system,'' Hearing Res. **60,** 115–142.
Liberman, M. C. (**1984**). ''Single-neuron labeling and chronic cochlear pathology. I. Threshold shift and characteristic-frequency shift,'' Hearing Res. **16,** 33–41.
May, B. J., Huang, A., LePrell, G., and Hienz, R. D. (**1996**). ''Vowel formant frequency discrimination in cats: Comparison of auditory nerve representations and psychophysical thresholds,'' Aud. Neurosci. **3**, 135–162
Meddis, R., and Hewitt, M. J. (**1992**). ''Modeling the identification of concurrent vowels with different fundamental frequencies,'' J. Acoust. Soc. Am. **91,** 233–245.
Miller, M. I., and Sachs, M. B. (**1984**). ''Representation of voice pitch in discharge patterns of auditory-nerve fibers,'' Hearing Res. **14,** 257–279.
Palmer, A. R. (**1990**). ''The representation of the spectra and fundamental frequencies of steady-state single- and double-vowel sounds in the temporal discharge patterns of guinea pig cochlear-nerve fibers,'' J. Acoust. Soc. Am. **88,** 1412–1426.
Palmer, A. R., Winter, I. M., and Darwin, C. J. (**1986**). ''The representation of steady-state vowel sounds in the temporal discharge patterns of the guinea pig cochlear nerve and primarylike cochlear nucleus neurons,'' J. Acoust. Soc. Am. **79,** 100–113.
Palmer, A. R., Winter, I. M., Gardner, R. B., and Darwin, C. J. (**1987**). ''Changes in the phonemic quality and neural representation of a vowel by alteration of the relative phase of harmonics near *F*1,'' in *The Psychophysics of Speech Perception*, edited by M. E. H. Schouten (Martinus Nijhoff, Dordrecht), pp. 371–376.
Parsons, T. W. (**1976**). ''Separation of speech from interfering speech by means of harmonic selection,'' J. Acoust. Soc. Am. **60,** 911–918.
Peterson, G. E., and Barney, H. L. (**1952**). ''Control methods used in a study of the vowels,'' J. Acoust. Soc. Am. **24,** 175–184.
Prijs, V. F. (**1989**). ''Lower boundaries of two-tone suppression regions in the guinea pig,'' Hearing Res. **42,** 73–82.
Rhode, W. S., and Greenberg, S. (**1992**). ''Physiology of the cochlear nucleus,'' in *The Mammalian Auditory Pathway: Neurophysiology*, edited by A. N. Popper and R. R. Fay (Springer-Verlag, New York), pp. 94–152.
Rhode, W. S., and Greenberg, S. (**1994**). ''Encoding of amplitude modulation in the cochlear nucleus of the cat,'' J. Neurophysiol. **71,** 1797–1825.
Sachs, M. B. (**1984**). ''Speech encoding in the auditory nerve,'' in *Hearing Science Recent Advances*, edited by C. I. Berlin (College-Hill, San Diego), pp. 263–308.
Sachs, M. B., and Young, E. D. (**1979**). ''Encoding of steady-state vowels in the auditory nerve: Representation in terms of discharge rate,'' J. Acoust. Soc. Am. **66,** 470–479.
Sachs, M. B., and Young, E. D. (**1980**). ''Effects of nonlinearities on speech encoding in the auditory nerve,'' J. Acoust. Soc. Am. **68,** 858–875.
Sachs, M. B., Voigt, H. F., and Young, E. D. (**1983**). ''Auditory nerve representation of vowels in background noise,'' J. Neurophysiol. **50,** 27–45.
Sachs, M. B., Winslow, R. L., and Blackburn, C. C. (**1988**). ''Representation of speech in the auditory periphery,'' in *Auditory Function: Neurobiological Bases of Hearing*, edited by G. E. Edelman, W. E. Gall, and W. M. Cowan (Wiley, New York), pp. 747–774.
Schreiner, C. E., and Langner, G. (**1988**). ''Periodicity coding in the inferior colliculus of the cat. II. Topographical organization,'' J. Neurophysiol. **60,** 1823–1839.
Scheffers, M. T. M. (**1983**). ''Sifting vowels: Auditory pitch analysis and sound segregation,'' Ph.D. thesis, Groningen Univ.
Sinex, D. G., and Geisler, C. D. (**1983**). ''Responses of auditory-nerve fibers to consonant-vowel syllables,'' J. Acoust. Soc. Am. **73,** 602–615.
Sokolich, W. G. (**1977**). ''Improved acoustic system for auditory research,'' J. Acoust. Soc. Am. Suppl. 1 **62,** S12.
Stubbs, R. J., and Summerfield, Q. (**1988**). ''Evaluation of two voice-separation algorithms using normal-hearing and hearing-impaired listeners,'' J. Acoust. Soc. Am. **84,** 1236–1249.
Wang, X., and Sachs, M. B. (**1993**). ''Neural encoding of single-formant stimuli in the cat. I. Responses of auditory nerve fibers,'' J. Neurophysiol. **70,** 1054–1075.
Wang, X., and Sachs, M. B. (**1994**). ''Neural encoding of single-formant stimuli in the cat. II. Responses of anteroventral cochlear nucleus units,'' J. Neurophysiol. **71,** 59–78.
Wang, X., and Sachs, M. B. (**1995**). ''Transformation of temporal discharge patterns in a ventral cochlear nucleus stellate cell model: Implications for physiological mechanisms,'' J. Neurophysiol. **73,** 1600–1616.
Weintraub, M. (**1987**). ''Sound separation and auditory perceptual organization,'' in *The Psychophysics of Speech Perception*, edited by M. E. H. Schouten (Martinus Nijhoff, Dordrecht), pp. 125–134.
Wiener, F. M., and Ross, D. A. (**1946**). ''The pressure distribution in the auditory canal in a progressive sound field,'' J. Acoust. Soc. Am. **18,** 401–408.
Winslow, R. L., Barta, P. E., and Sachs, M. B. (**1987**). ''Rate coding in the

auditory nerve,'' in *Auditory Processing of Complex Sounds*, edited by C. S. Watson and W. Yost (Erlbaum, Hillsdale, NJ), pp. 212–224

Winter, I. M., and Palmer, A. R. (**1990a**). ''Responses of single units in the anteroventral cochlear nucleus of the guinea pig,'' Hearing Res. **44,** 161–178.

Winter, I. M., and Palmer, A. R. (**1990b**). ''Temporal responses of primary-like anteroventral cochlear nucleus units to the steady-state vowel /i/,'' J. Acoust. Soc. Am. **88,** 1437–1441.

Young, E. D., and Sachs, M. B. (**1979**). ''Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory nerve fibers,'' J. Acoust. Soc. Am. **66,** 1381–1403.

Young, E. D., Robert, J.-M., and Shofner, W. P. (**1988**). ''Regularity and latency of units in the ventral cochlear nucleus: implications for unit classification and generation of response properties,'' J. Neurophysiol. **60,** 1–29.

Two separate codes for missing-fundamental pitch in the human auditory cortex

I. Winkler[a)]
Department of Psychophysiology, Institute for Psychology, Hungarian Academy of Sciences, P.O. Box 398, Szondi u. 83/85, H-1394 Budapest, Hungary and Cognitive Brain Research Unit, Department of Psychology, University of Helsinki, P.O. Box 13, FIN-00014 University of Helsinki, Finland

M. Tervaniemi and R. Näätänen
Cognitive Brain Research Unit, Department of Psychology, University of Helsinki, P.O. Box 13, FIN-00014 University of Helsinki, Finland

(Received 19 November 1996; revised 3 March 1997; accepted 20 April 1997)

Two auditory event-related potential components, the supratemporal N1 and the mismatch negativity (MMN), index traces encoding the missing-fundamental pitch. The present results suggest that these two codes derive from separate pitch extraction processes. Frequent 300-Hz and infrequent 600-Hz missing-fundamental tones were presented, in some stimulus blocks with short (150 ms), in others, with long (500 ms) stimulus durations. MMN, reflecting a preattentive change detection process, was elicited by infrequent missing-fundamental tones only in the long-duration condition. Correspondingly, subjects were able to detect these high-pitch missing-fundamental tones amongst similar low-pitch ones only when the stimulus duration was long. In addition, the MMN response peaked ca. 120 ms later for missing-fundamental tones than for pure tones of the fundamental frequency suggesting that missing-fundamental pitch resolving took substantially longer than extracting the spectral pitch. In contrast, a differential N1 response to the missing-fundamental pitch was found for both stimulus durations, with no substantial difference in peak latency between the pure and missing-fundamental tones. The contrasting features found for the two auditory cortical missing-fundamental pitch codes support the notion of two separate missing-fundamental pitch resolving mechanisms. [S0001-4966(97)05208-9]

PACS numbers: 43.64.Ri, 43.66.Fe, 43.66.Hg [RDF]

INTRODUCTION

The perceived pitch of complex tones consisting of several partials corresponds to that of the fundamental whether or not the fundamental itself is present in the sound. The utility of resolving the missing-fundamental pitch in natural environments is stressed by the presence of this ability in animals (e.g., Tomlinson and Schwarz, 1988) and by its early appearance in the course of human ontogenesis (Clarkson and Clifton, 1985). The importance of the missing-fundamental pitch phenomenon (Seebeck, 1841; for recent reviews, see de Boer, 1976; Scharf and Houtsma, 1986) for pitch perception theories lies in the fact that it demonstrates that pitch analysis extends beyond the auditory periphery as the pitch experience for complex tones does not fully correspond to what could be expected on the basis of the peripheral distortion (Schouten *et al.*, 1962). In addition, pitch recognition for missing-fundamental tones does not deteriorate when harmonics are presented dichotically (Houtsma and Goldstein, 1972). Therefore recent pitch perception theories suggest that central pattern-recognition processes extract the fundamental frequency from the aurally resolved low-order harmonics (Goldstein, 1973; Wightman, 1973; Terhardt, 1974).

Studies showing how pitch information is encoded in the first stages of the afferent auditory pathway provided data for determining the constraints of the subsequent processing of the missing-fundamental pitch. However, signals measured from 8th-nerve fibers (e.g., Horst *et al.*, 1986), single cells of the cochlear nucleus (e.g., Evans, 1977), or auditory evoked brainstem potentials (Galbraith, 1994) are only indirectly related to the perceived pitch of missing-fundamental tones. Cortically generated auditory event-related potentials (ERP) may provide an important link between neural activity and pitch perception.

ERP responses can be decomposed into contributions of distinct generator processes, the ERP components, appearing in their typical latency range and displaying the scalp distribution characteristic to the underlying generator (Näätänen and Picton, 1987). Some of these components are associated with perceptual processes. One piece of such evidence was provided by Pantev *et al.* (1989), who found that the generator location of the magnetic counterpart of the auditory N1 potential (N1m) elicited by complex missing-fundamental tones corresponded to the auditory cortical source activated by pure tones of the fundamental frequency. Pantev *et al.* (1996) extended the previous results to dichotically presented missing-fundamental tones: the source location of the N1m component elicited when odd and even harmonics of the same fundamental frequency were presented to different ears matched that of the N1m to a pure tone of the fundamental frequency.

[a)]Electronic mail: winkler@cogpsyphy.hu

The auditory N1 wave is a central negativity typically peaking between 80 and 140 ms from stimulus onset (for a review, see Näätänen and Picton, 1987). The N1m predominantly reflects the contribution of the auditory cortex to the electric N1 wave. This N1 constituent was termed the supratemporal N1 component (cf. Vaughan and Ritter, 1970; for a detailed discussion of further N1 subcomponents, see Näätänen and Picton, 1987). The neural elements generating the N1m are tonotopically organized in the primary auditory cortex, that is, the stimulus frequency is encoded by the location of the evoked neural activity (Elberling *et al.*, 1982; Romani *et al.*, 1982; Pantev *et al.*, 1988). Frequency (pitch) traces are probably preserved by the supratemporal N1 neurons in the form of refractoriness patterns demonstrated by the frequency-specific decrease of the N1 (N1m) amplitude to sounds of similar frequency presented in short succession (Hari *et al.*, 1982; Näätänen *et al.*, 1988; Sams and Salmelin, 1994). This response decrement cannot be reversed by dishabituation procedures (Ritter *et al.*, 1968; Barry *et al.*, 1992). Therefore, Pantev *et al.*'s (1989, 1996) results suggest that a missing-fundamental pitch code is present in the neuronal population generating the supratemporal N1 by ca. 80 ms from the stimulus onset (the time the N1m response reflecting the missing-fundamental pitch commences).

However, there exists a large body of evidence showing that the N1 response does not correspond to auditory perception (for reviews, see Näätänen and Picton, 1987; Näätänen, 1992). For example, in terms of the N1 wave, a 1000-Hz tone amplitude modulated at 200 Hz is more similar to a 1000-Hz pure tone than a 200-Hz one (Butler, 1972), whereas the perceived pitch of this complex tone is approximately equal to that of a 200-Hz tone. In addition, the N1 amplitude does not correlate with pitch recognition (Parasuraman *et al.*, 1982).

In a recent study, Winkler *et al.* (1995) found that infrequent complex missing-fundamental tones of 600-Hz fundamental frequency presented amongst a set of similar tones having 300 Hz as the missing fundamental elicited an MMNm response, the magnetic counterpart of the mismatch negativity (MMN) ERP component. The MMN is a frontocentrally negative potential evoked by infrequent changes (deviant stimuli) in a sequence of a repetitive (standard) auditory stimulus or stimulus feature (for recent reviews, see Näätänen, 1992; Ritter *et al.*, 1995). Because the MMN is not elicited by frequently presented stimuli or by a rare stimulus per se in the absence of a frequent stimulus (Lounasmaa *et al.*, 1989; Näätänen *et al.*, 1989), it appears that MMN reflects the deviation of the incoming stimulus from the auditory sensory memory trace of the frequently presented stimulus or stimulus feature (Näätänen *et al.*, 1978; Näätänen, 1992; Cowan, 1995). The accuracy of the sound representations indexed by the MMN seems to correspond to the perceptual resolution of auditory stimulus features (see Näätänen and Alho, 1995; Kraus *et al.*, 1996). MMN is elicited even when subjects perform some task unrelated to the auditory stimulation (like reading a book or performing some visual task; e.g., Lyytinen *et al.*, 1992). Therefore the change detection process reflected by MMN is probably preattentive. Thus MMN provides accurate information about the underlying memory representations, which is free from task-related distortions.

The composition and within-block probabilities of Winkler *et al.*'s (1995) missing-fundamental tones was such that each complex tone, as well as each harmonic frequency was presented equiprobably in the stimulus blocks (the virtual pitch conditions of the present experiments I and II use the same design, see below). This way, the occasional change from 300- to 600-Hz virtual pitch was not accompanied by similarly infrequent changes in spectral pitch. The MMN elicited by the 600-Hz missing-fundamental tone with respect to the 300-Hz ones demonstrated that the memory traces involved in the process generating this component represented the virtual pitch of complex tones, not the spectral pitch of their harmonic components. Infrequent deviations in periodicity pitch, another typical way of separating spectral and virtual pitch, also elicit an MMN (Lu *et al.*, 1989).

Winkler *et al.*'s (1995) MMNm to virtual-pitch deviants originated from the auditory cortex, in accordance with the well-known auditory cortical location of the primary MMN generator (Hari *et al.*, 1984; Halgren *et al.*, 1995; for a review, see Alho, 1995). The peak latency of this pitch-MMNm was, however, longer by ca. 120 ms (200–280 ms) than the typical peak latency (100–140 ms) for comparable changes in the spectral pitch of pure tones (see Näätänen, 1992). This result suggested that resolving the missing-fundamental pitch took substantially longer than pitch extraction from pure tones. On the other hand, the orientation of the N1m generator activated by the missing-fundamental deviant stimulus differed from that of the N1m generator for the standard stimuli. This latter result showed that the missing-fundamental pitch information was also encoded by the neuronal population underlying the N1m response, as was demonstrated by Pantev *et al.* (1989, 1996).

If the missing-fundamental pitch code was complete at ca. 80 ms after stimulus onset, judging from the onset of the N1m component, why then did preattentive detection of deviance take 120 ms longer for missing fundamental than for pure tones (the latter elicit N1ms of similar onset latencies)? Two alternative explanations can be considered.

(1) Independent codes hypothesis: The missing-fundamental pitch information reflected by the N1m and MMNm are determined by two separate pitch-resolving processes of different durations.

(2) Common origin hypothesis: The pitch information reflected by the N1m and MMNm derives from that same pitch-resolving process. To account for the delayed MMNm to missing-fundamental tones compared with that to pure tones (Winkler *et al.*, 1995; whereas no such difference was observed for the N1m; Pantev *et al.*, 1989, 1996), one should assume that the original pitch code undergoes further processing before it is stored in the stimulus representations indexed by the MMNm.[1] This latter process takes longer for virtual than for spectral pitch.

The present study investigated the relationship between the above two auditory cortical codes of missing-fundamental pitch. The independent codes hypothesis suggests that the traces reflected by the supratemporal N1 component are established by a fast process capable of

TABLE I. Partials (in Hz) used in composing the complex missing fundamental tones. The frequencies employed in creating the deviant (600 Hz) tone are italicized throughout the table.

Missing fundamental component	Standards 300 Hz									Deviant 600 Hz
	1	2	3	4	5	6	7	8	9	
1	*1200*	*1200*	*1200*	*1200*	*1200*	1500	1500	1500	*1800*	*1200*
2	1500	1500	1500	*1800*	2100	*1800*	*1800*	2100	2100	*1800*
3	*1800*	2100	*2400*	2100	*2400*	2100	*2400*	*2400*	*2400*	*2400*

determining the missing-fundamental pitch from short acoustic segments. In contrast, the auditory stimulus representations indexed by the MMN component result from a slower process, one, perhaps, requiring longer acoustic samples. To separate the two assumed pitch-resolving processes, the present set of experiments investigated the effects of stimulus duration on the pitch discrimination of missing-fundamental tones and on the target ERP responses, the N1 and MMN. As the N1 peak latency does not increase with increasing stimulus durations, the process encoding the missing-fundamental pitch for N1 should not be affected by stimulus durations exceeding the well-known range of the N1 peak latency. Therefore if by employing different stimulus durations (all longer than the N1 latency) the elicitation or the latency of the MMN response can be influenced, one should conclude that the pitch-resolving process underlying the auditory stimulus representations of the MMN system proceeds separately from the one whose outcome is reflected by the N1. On the other hand, if such stimulus durations did not produce differential MMN effects, the more parsimonious explanation of common pitch resolving can be maintained. In addition, it was expected that voluntary discrimination performance would parallel the emergence of the pitch code reflected by the MMN as previous investigations suggested good correspondence between MMN and voluntary discrimination performance (for a review, see Näätänen and Alho, 1995).

I. EXPERIMENTS I AND II

A. Procedure

Nine subjects with normal audiological status (five females, 18–29 years, mean age 23.2 years) participated in experiment I, 12 subjects (eight females, 19–27 years, mean age 22 years, all different from those of experiment I) in experiment II. Subjects gave informed consent after the nature of the study was explained to them. During the EEG recordings, the subject was sitting in a reclining chair in an acoustically dampened and electrically shielded room reading a book of his/her choice.

Six blocks of 500 short tones were binaurally delivered to subjects via TDH-50p headphones (NeuroStim stimulation system) at an intensity level of 75 dB, sound-pressure level. Half of the blocks contained randomized sequences of 300-Hz (''low,'' 90%) and 600-Hz (''high,'' 10%) pure tones (spectral pitch condition). In the other three stimulus blocks, ten different complex missing-fundamental tones were equiprobably delivered at 10% probability (virtual pitch condition). Nine out of the ten complex tones had its missing-fundamental frequency at 300 Hz, the tenth at 600 Hz. The missing-fundamental tones were composed of three partials selected from the range between the 3rd and the 7th overtones of the 300-Hz fundamental (equaling the 1st to 3rd overtones of the 600-Hz fundamental). Table I gives the frequencies used for constructing the complex tones. Each individual harmonic frequency was employed in six out of the ten missing-fundamental tones (including the 600-Hz one). In addition, pink noise covering the range of the two fundamentals (200–800 Hz) was added to each complex tone at an intensity level of -40 dB relative to the intensity of the summed partials to prevent the fundamentals from emerging in the output signal due to nonlinear distortions of the stimulation equipment. The overall intensity was identical to that in the spectral pitch condition. The order of the stimulus blocks was balanced across subjects.

In experiment I (''short duration''), tones of 150-ms duration were delivered with an interstimulus interval (offset-to-onset; ISI) of 350 ms. In experiment II (''long duration''), 500-ms tones were presented with 400 ms ISI. Stimulus durations included 5-ms rise and 5-ms fall times.

The electroencephalogram (EEG) was measured with Ag/AgCl electrodes from Fpz, Fz, Cz, Pz, and Oz (10-20 system), both mastoids (LM and RM, for the left and right mastoids, respectively), and from L1, L2, R1, and R2 (the one- and two-third points on the arc connecting Fz with the mastoids on both sides). The horizontal electrooculogram (HEOG) was monitored at the outer canthus of the right eye. The common reference electrode was placed on the tip of the nose. The EEG and EOG signals were digitized at a sampling rate of 500 Hz with bandpass filtering between 0.1 and 100 Hz (-3 dB points). Responses (analysis period: 400 ms starting 50 ms before stimulus onset) were filtered (low-pass bandlimit 30 Hz) and separately averaged for each stimulus type. Epochs containing an EEG or EOG change exceeding 100 μV at any lead were omitted from the analysis. Amplitude measurements were referred to the mean amplitude of the prestimulus period (''biological zero''). Statistical analyses, t tests and dependent analyses of variance (ANOVA), were conducted by the BMDP statistical program package (Dixon, 1985).

B. Results

In the spectral pitch condition of both experiments, infrequent high tones (deviants) were preattentively detected amongst repetitive sequences of the low tones. This was demonstrated by the larger negativity elicited in response to deviant than standard stimuli in the 100–150-ms interval

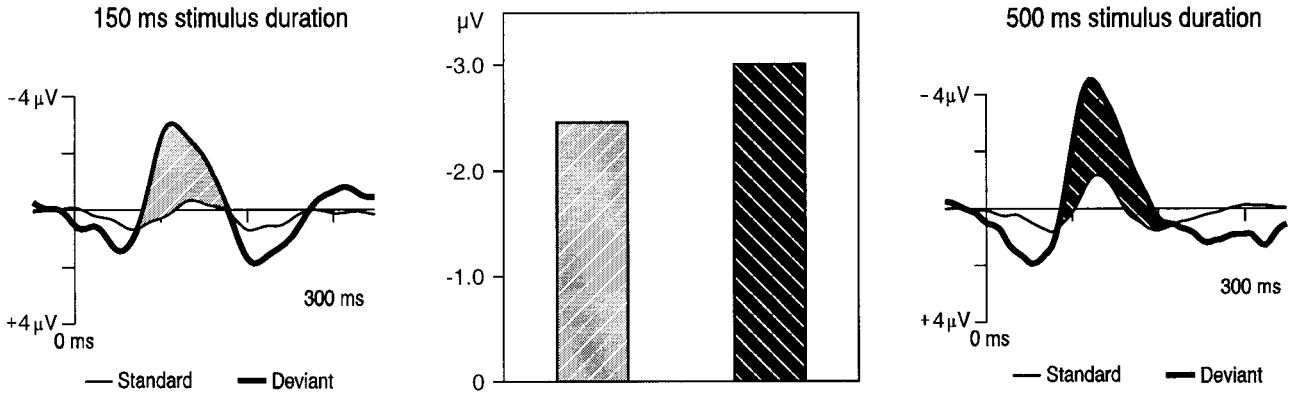


FIG. 1. Experiments I and II, spectral pitch condition: Frontal (Fz) grand-average ERP responses to frequent low (standard, thin line) and infrequent high (deviant, thick line) pure tones in the short (left side) and long (right side) stimulus-duration experiments. The negative difference between the deviant and standard responses (deviant-standard) is marked for each stimulus duration. The bar chart at the center of the figure gives the amplitude of the correspondingly marked difference responses.

from stimulus onset (Fig. 1, see Figs. 3 and 4, left side for individual responses in the spectral pitch conditions). The statistical significance of the mean frontal (Fz) difference between the standard and deviant responses in the 100–150-ms interval was verified by 1-group t tests [$t(8)=5.09$, mean difference: $d=2.46\ \mu$V, Standard Error of Mean: SEM$=0.48$, $p<0.001$ and $t(11)=7.3$, $d=-3.01\ \mu$V, SEM$=0.41$, $p<0.0001$, for the short- and long-duration stimuli, respectively]. The large extra negativity in response to deviant stimuli (compared to that to the standards) consisted predominantly of the MMN component which was only elicited by infrequent stimuli. The scalp distribution of this additional negativity had a frontocentral maximum, was slightly larger over the right than left hemisphere appearing with reversed polarity at the mastoid. This is compatible with the notion that MMN was elicited by the deviant stimuli (see, Näätänen, 1992 for a description of the MMN scalp distribution). With large (one octave) separation between the standard and deviant frequencies, MMN is known to overlap the N1 wave (Scherg *et al.*, 1989). Therefore, part of the difference between the standard and deviant responses was probably due to frequency-specific refractoriness differences in the N1 components. In other words, the N1 neurons responding to the 300-Hz stimulus frequency could have become more refractory during the frequent repetition of the 300-Hz stimulus than the neurons activated only by the 600-Hz tone. The central (Cz) N1 peak latency in the grand-average responses to low standard tones was 132 and 124 ms for long and short tones, respectively (the overlap between the N1 and MMN prevented the measurement of the N1 peak latencies for the deviant tones).

Infrequent high-pitch missing-fundamental tones amongst frequent low-pitch ones were preattentively detected only at the long stimulus duration (experiment II). Figure 2 (right side) shows the MMN component at the 200–260-ms interval of the frontal responses to infrequent high-pitch missing-fundamental tones [$t(11)=4.59$, $d=-0.48\ \mu$V, SEM$=0.098$, $p<0.001$, for the mean difference in this interval], whereas the tiny negative difference appearing in the same interval for the short tones (Fig. 2, left side) was far from being significant [$t(8)=0.97$, $d=0.053\ \mu$V, SEM $=0.054$]. Individual responses (Figs. 3 and 4, right side) demonstrate that the two experiments yielded separate, but consistent patterns of results with the expected interindividual variance. In experiment II, the scalp distribution of the difference between the deviant and standard responses in the 200–260-ms interval was supporting the identification of this negativity as an MMN component [see also Winkler *et al.* (1995), where the equivalent current dipole of the corresponding magnetic response was localized in the auditory cortex].

In contrast, with both stimulus durations, low-pitch missing-fundamental standards elicited a later and somewhat larger N1 component than the high-pitch deviant stimulus (Fig. 2). The mean frontal (Fz) amplitudes of the standard and deviant responses differed in the 120–170-ms interval [$t(8)=2.44$, $d=0.37\ \mu$V, SEM$=0.15$, $p<0.05$ and $t(11)=3.66$, $d=0.65\ \mu$V, SEM$=0.18$, $p<0.01$, for experiments I and II, respectively]. More importantly, the peak latency of the central (Cz) N1 response (measured separately for each subject within the 80–140-ms interval) was significantly longer for the low-pitch than for the high-pitch missing-fundamental tones [$t(8)=3.19$, $d=10.2\ \mu$V, SEM$=3.19$, $p<0.05$, and $t(11)=3.28$, $d=9.0$ ms, SEM$=2.75$, $p<0.01$, for short and long stimulus durations, respectively]. The central (Cz) N1 peak latency in the grand-average responses was 128 ms for the low-pitch standards of both stimulus durations, 120 and 108 ms for the short- and long-duration high-pitch deviants, respectively.

In general, the N1 amplitudes measured in experiment II were higher than those in experiment I. This was due to the difference in stimulus onset asynchronies (onset-to-onset interval; SOA being 500 vs 900 ms, in experiments I and II, respectively). The N1 amplitude increases with increasing SOAs below ca. 10 s as a consequence of the long recovery period of the N1-generating neuronal elements (Hari *et al.*, 1982; for a review see Näätänen and Picton, 1987).

The MMN (MMN+N1) responses shown in Figs. 1 and 2 were usually followed by an increased positivity in the 200–350-ms latency range (see, also Figs. 3 and 4). The frontocentral scalp distribution of this wave suggests that in

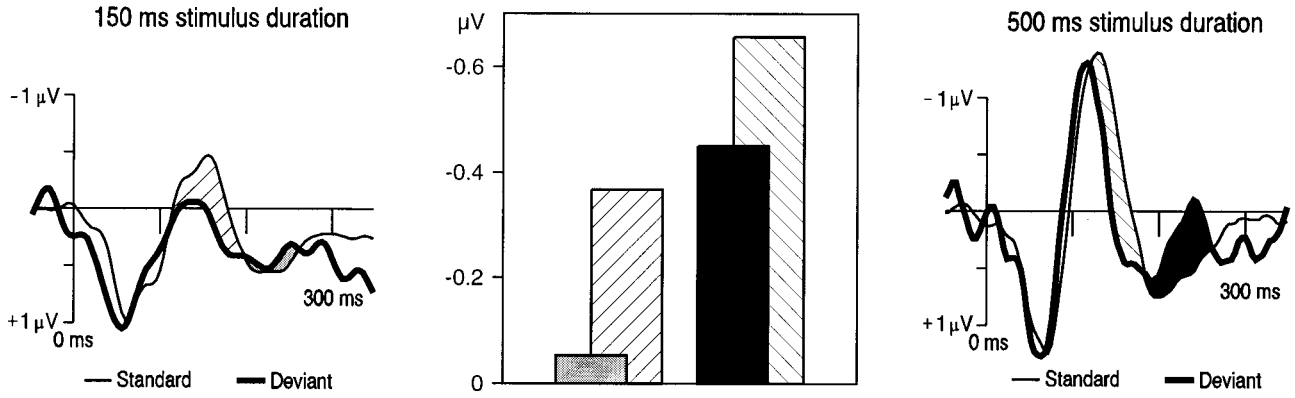


FIG. 2. Experiments I and II, virtual pitch condition: Frontal (Fz) grand-average ERP responses to frequent low- (standard, thin line) and infrequent high-pitch (deviant, thick line) missing-fundamental tones in the short (left side) and long (right side) stimulus-duration experiments. The N1 difference between the deviant and standard responses (hatched patterns; standard-deviant) and the MMN component (solid shades; deviant-standard) are marked for each stimulus duration. The bar chart at the center of the figure gives the amplitude of the correspondingly marked difference responses.

some subjects, deviant stimuli might have elicited the P3a component (Squires *et al.*, 1975) which often follows the MMN response (Lyytinen *et al.*, 1992). The P3a is assumed to accompany attention switching, a possible outcome of preattentive change detection (cf. Näätänen, 1992).

C. Discussion

Experiments I and II established that the process reflected by the MMN component requires long (>150 ms) acoustic samples of the complex tones to preattentively detect infrequent high-pitch missing-fundamental tones embedded in sequences of similar low-pitch tones. For pure tones of the equivalent spectral pitch, stimuli of 150-ms duration proved to be sufficient. In addition, the MMN to missing-fundamental pitch change peaked ca. 120 ms later than that elicited by an equal amount of deviation in spectral pitch. These results suggest that resolving and encoding the virtual pitch of missing-fundamental tones for the auditory stimulus representations underlying preattentive change detection takes significantly longer than the processing of spectral pitch. The alternative explanation that the late MMN to missing-fundamental pitch deviants (compared to that to pure-tone ones) was the result of prolonged change detection (rather than of slower pitch resolving for missing fundamental than pure tones) is contrasted by the lack of MMN to short-duration missing-fundamental pitch deviants. If the MMN latency difference between the long-duration spectral pitch (pure tone) and virtual pitch (missing fundamental) conditions were caused by longer preattentive change detection for missing fundamental than for spectral pitch, short-duration missing-fundamental pitch deviants should also have elicited the MMN (if only later than the short-duration pure-tone deviants).

In contrast to the MMN, the N1 response to the two missing-fundamental tones differed at both stimulus durations, with the latency of the N1 wave being approximately equal in the responses to pure and missing-fundamental tones.

Experiment III was designed to test whether the observed differences in the N1 responses to low- and high-pitch missing-fundamental tones reflected the pitch-specific nature of the N1 wave or whether this difference was rather caused by the unequal probabilities of stimulus presentation (by producing different refractoriness states in the N1 generating neurons responding to 300- versus 600-Hz pitch). To eliminate the effects of the unequal stimulus probabilities, each of the ten different missing-fundamental tones were presented in separate (homogeneous) blocks.

II. EXPERIMENT III

A. Procedure

Seven subjects with normal audiological status (four females, 19–25 years, mean age 21.6 years) participated in the experiment. Subjects gave informed consent after the nature of the study was explained to them. The stimuli, procedure, EEG/EOG recording, and measurements were identical to those of the virtual pitch condition of experiment II (500-ms stimulus duration), except that all ten different missing-fundamental tones were presented in separate blocks, each carrying 120 identical stimuli.

B. Results

Figure 5 shows that the N1 response peaked later for low-pitch than for high-pitch missing-fundamental tones causing an amplitude difference on the downward slope of this wave, similar to that observed in experiments I and II. The N1 peak-latency difference was comparable to that in experiments I and II: the grand-average peak latencies were 110, 106, 104, 104, 104, 112, 108, 106, and 106 ms for the nine low-pitch complex tones (the order corresponds to that in Table I), 98 ms for the high-pitch missing-fundamental tone (the difference between the N1 peak latencies in response to low- and high-pitch tones was ca. 8 ms, on the average). An analysis of variance of all ten responses (one factor with ten levels for the ten N1 peak-latency measurements conducted separately for each subject within the 80–140-ms interval) confirmed the observed N1 peak latency effect [$F(9,54)=3.85$, $p<0.01$, the Huynh–Feldt adjustment factor for the degrees of freedom being 0.879]. The *post-hoc* contrast showed that the N1 peak latency to the high-pitch complex tone was significantly shorter than that to the low-

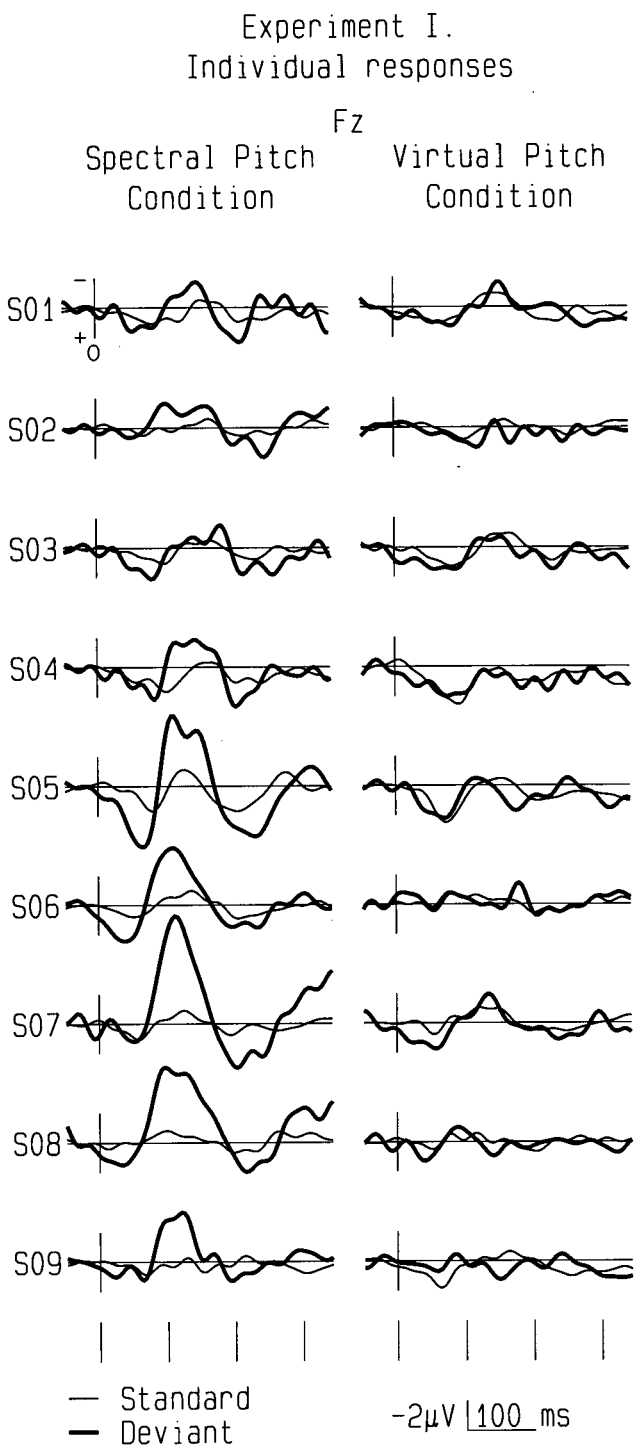


FIG. 3. Experiment I: Frontal (Fz) responses to frequent low- (standard, thin line) and infrequent high-pitch pure (left side) and missing-fundamental tones (right side) separately for each subject. Subjects are identified by the codes S01 through S09.

pitch ones [$F(1,6)=11.83$, $p<0.02$]. The amplitude difference (higher mean N1 amplitudes for the low complex tones) in the 120–170-ms range (the downward slope of the N1, early P2) was also significant [$t(6)=3.15$, $d=0.69\ \mu$V, SEM$=0.22$, $p<0.05$].

C. Discussion

The present results demonstrated that the neuronal elements generating the N1 response differentially responded to

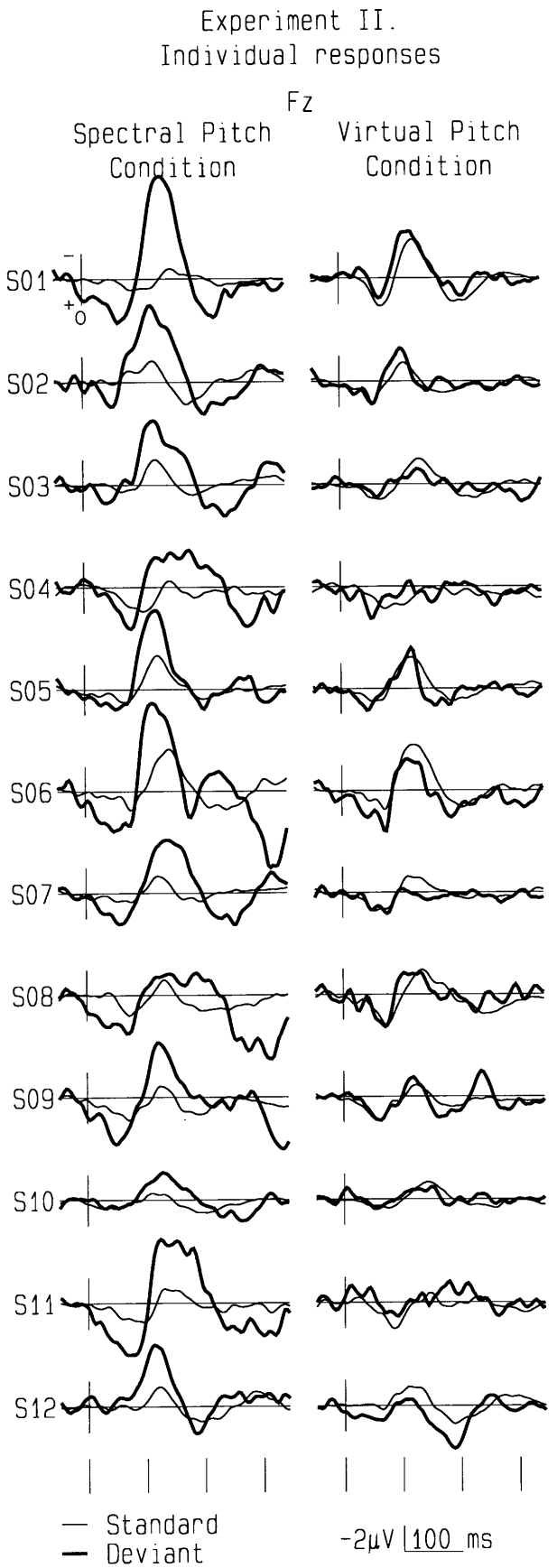


FIG. 4. Experiment II: Frontal (Fz) responses to frequent low- (standard, thin line) and infrequent high-pitch pure (left side) and missing-fundamental tones (right side) separately for each subject. Subjects are identified by the codes S01–S12 (all being different from those participating in experiment I).

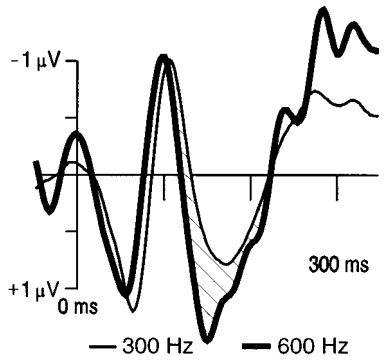


FIG. 5. Experiment III: Frontal (Fz) grand-average responses to low- and high-pitch missing fundamental tones presented in separate stimulus blocks. Responses to different low-pitch missing fundamental tones were collapsed. The N1 (and the consequent P2) difference between the low- and high-pitch tones is marked in the figure by hatched pattern fill.

the pitch of missing-fundamental tones. The N1 peak latency of all nine 300-Hz pitch missing fundamental tones was longer than that of the 600-Hz pitch one. The pitch specificity of the N1 response did not depend on the within-block probability of stimulus presentation as, in this experiment, each tone was presented alone (i.e., without mixing it with other tones within the stimulus blocks). A reduction of the N1 peak latency with increased missing-fundamental pitch (100–330 Hz) has also been observed by Ragot and Lepaul-Ercole (1996). These authors employed short (200 ms) stimulus durations. In the present study, the N1 peak latency was not affected by the spectral pitch of the complex tone components. This was demonstrated by the lack of systematic variation in the pattern of N1 peak latencies (comparing the spectral composition of the 300-Hz pitch complex tones given by Table I with the grand-average N1 peak latencies measured from the corresponding ERP responses, see Results above). If the N1 peak latency was, at least partially, determined by the spectral pitch of the tonal components, one should expect the similarity between the N1 peak latencies responding to two complex tones of identical virtual pitch to depend on the number of common frequency components.

The virtual pitch-specific nature of the N1 neurons is demonstrated by the latency effect in the electrically recordable N1 wave and by the pitch-dependent generator location of the N1m (the magnetic N1) component (Pantev *et al.*, 1989, 1996; Winkler *et al.*, 1995). The fact that also the peak latency of the N1/N1m elicited by pure tones decreases with increasing tone frequencies below 1000 Hz (Jacobson *et al.*, 1992; Roberts and Poeppel, 1996) supports Pantev *et al.*'s (1989) conclusion that the N1m neurons responding to spectral pitch might also be involved in the N1m responses to tones of the corresponding missing-fundamental pitch.

Thus the present results support the notion that the neuronal population generating the supratemporal N1 component encode the pitch of missing-fundamental stimuli. The results of experiment I and those of Ragot and Lepaul-Ercole (1996) demonstrated that the missing-fundamental pitch is encoded by the auditory cortical N1 neurons even at short stimulus durations.

Experiment IV tested the effects of stimulus duration on the voluntary detection of infrequent high-pitch pure or missing-fundamental tones amongst similar frequent low-pitch tones.

III. EXPERIMENT IV

A. Procedure

Twelve subjects of normal audiological status (five females, 20–27 years, mean age 22.3 years) participated in the experiment. Subjects were instructed to press a reaction button when detecting a high-pitch tone. Four stimulus blocks, each corresponding to one of the spectral pitch (pure tone) or virtual pitch (missing-fundamental tone) conditions of experiments I (150-ms tone duration) or II (500-ms tone duration), were presented to subjects. Each stimulus block was preceded by a short training session. Stimulation parameters were identical to those of the corresponding experiment/condition, except that the present blocks consisted of 250 stimuli. The order of the stimulus blocks was balanced across subjects.

B. Results

Due to the exact octave difference in pitch, the required discrimination of missing-fundamental stimuli was difficult for most subjects (cf. Krumhansel, 1990). With the short tone duration, only a few of them managed to detect the high-pitch missing-fundamental tones above the chance level (mean detection rate 19.33%, with 10% of the tones having high missing-fundamental pitch). However, the discrimination performance for missing-fundamental tones dramatically improved with the long stimulus duration (mean detection rate 62.5%). The discrimination of pure tones was close to perfect at both stimulus durations. Figure 6 shows the group-average Grier A′ sensitivity values (Grier, 1971; Aaronson and Walls, 1987) calculated separately for each stimulus block. An ANOVA analysis of these scores [tone type (pure versus missing fundamental)×tone duration (150 versus 500 ms)] revealed an interaction between the two factors [$F(1,11)=38.85$, $p<0.0001$]. This interaction was caused by the pattern of results showing significant performance difference between the two tone durations for missing fundamental [$F(1,11)=31.56$, $p<0.001$, $A'(150)=0.63$, $A'(500)=0.84$] but none for pure tones [$F(1,11)=1.62$, $A'(150)=0.99$, $A'(500)=0.95$]. There was a main effect of the tone-type factor confirming that it was easier to discriminate pure than missing-fundamental tones [$F(1,11)=55.15$, $p<0.0001$]. The corresponding bias (Grier's B″) values were higher for pure tones than for missing-fundamental tones and for long tones than for short tones [$F(1,11)=88.08$, $p<0.0001$ and $F(1,11)=11.21$, $p<0.01$, for tone type and tone duration, respectively]. No interaction was found between the two factors for the bias values. Reaction times could only be analyzed for the long tones as the number of detected short missing-fundamental targets was too small for several subjects. The average reaction time (RT) for long missing-fundamental tones was significantly longer than that

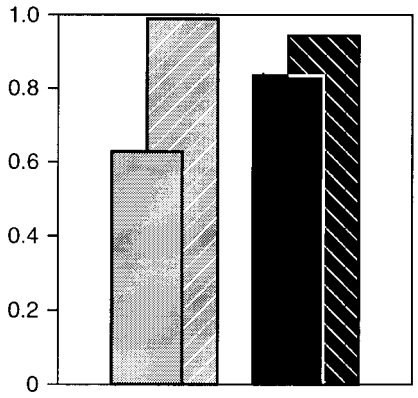


FIG. 6. Grier's A′ sensitivity values (grand mean) describing the performance in discriminating low and high pure tones (hatched patterns) and missing-fundamental tones (solid shades) for the short (left side, grey shade) and long (right side, black color) stimulus-duration conditions.

for the corresponding pure tones (518.12 and 377.08 ms, respectively; $t(11)=6.83$, $d=141.03$ ms, SEM$=20.66$, $p<0.001$).

C. Discussion

The present results demonstrated that short samples of the stimuli were insufficient for a reliable discrimination of the low and high missing-fundamental sounds even though pure tones of the corresponding pitches could be easily distinguished with this duration. With the long stimulus duration, however, also the missing-fundamental tones of one or the other pitch could be discriminated from each other. In accordance with the difficulty of the task, RTs were longer for missing-fundamental tones than for pure tones (this was tested only for the long stimuli).

IV. GENERAL DISCUSSION

The picture emerging from the present experiments has three salient features: (1) The preattentive change detection (revealed by the elicitation of the MMN component) parallels voluntary discrimination performance. High pure tones could be discriminated from low ones and also the MMN component was elicited with both stimulus durations. In contrast, subjects were able to detect the infrequent high-pitch missing-fundamental tones amongst frequent low-pitch ones only with the long stimulus duration. Accordingly, the MMN component only emerged at this stimulus duration, but not when the short stimulus duration was used. (2) The neuronal population generating the supratemporal N1 component encoded the pitch of the pure and missing-fundamental tones similarly (probably in the form of tonotopy, i.e., by following the place principle). This code distinguished between the two employed pitches for the missing-fundamental tones at both stimulus durations, even though subjects were unable to make the required voluntary discrimination with the short stimulus duration. (3) The onset latency of the N1 response suggests that the pitch code reflected by this component must be complete by ca. 80 ms (for both pure and missing-fundamental tones). Spectral pitch information was accessible by the preattentive change detection at about the same time (judging from the early onset of the MMNs in the spectral pitch conditions of experiments I and II). However, the preattentive detection of the missing-fundamental pitch deviants (emerging only with the long stimulus duration) commenced ca. 120 ms later than that for the corresponding pure tones and the average RT measured when subjects discriminated these complex tones was ca. 140 ms longer than the RT for the pure tones.

The above-described pattern of results revealed that the traces reflected by the supratemporal N1 and the MMN components encoded the outcome of two separate missing-fundamental pitch-resolving processes. As was expected, prolonging the missing-fundamental stimuli past the typical N1 latency range did not increase the N1 peak latency. However, MMN was only elicited with the long, but not with the short stimulus duration. If the missing-fundamental pitch code underlying the two components derived from a common pitch-resolving process, the MMN should have been elicited with both stimulus durations. This is because the differential N1 response to missing-fundamental pitch at the short stimulus duration demonstrated that a 150-ms sound segment provided sufficient spectral information for the corresponding pitch-resolving process. If the missing-fundamental pitch information was accessible by the preattentive change detection process at the short stimulus duration, MMN should have been elicited, just as it was for pure-tone deviants. Therefore the missing-fundamental pitch code involved in preattentive change detection (MMN) and the one reflected by the N1 could not have been determined by a common pitch-resolving process.

The present conclusion of the existence of two separate pitch-resolving processes might be somewhat surprising. However, redundant mechanisms have already been proposed to account for the missing-fundamental pitch phenomenon. Due to the interference between unresolved higher-order partials of a given fundamental frequency at the auditory periphery (the residue), in most situations, the envelope of the cochlear output signal carries sufficient information for detecting the frequency of the corresponding fundamental (e.g., Horst *et al.*, 1986). Thus the auditory system could determine the fundamental pitch on the basis of this residue (Schouten, 1949, 1970). It has been suggested (de Boer, 1976; Houtsma and Smurzinsky, 1990; Carlyon and Shackleton, 1994) that beside the central pattern-recognition processes mentioned in the Introduction a residue-based mechanism might also be involved in perceiving the pitch of missing-fundamental tones. The two pitch codes found in the present study, however, cannot be regarded as resulting from the above-mentioned two mechanisms (pattern recognition and residue), because both of the present missing-fundamental pitch codes were observed for tones composed of low-order harmonics.

The correspondence between the elicitation of MMN and detection performance suggests that the voluntary discrimination of missing fundamentals (and probably also perception) relies on the same pitch code as the preattentive change detection. A similar correspondence between the MMN and discrimination performance has been observed in auditory recognition masking (Winkler *et al.*, 1993). In ad-

dition, the MMN peak latency strongly correlates with the reaction time in discrimination tasks and, further, it provides the earliest ERP index to predict the latency of later, task-dependent ERP components (such as the N2b and P3; e.g., Novak *et al.*, 1992; Tiitinen *et al.*, 1994). Corroborating evidence was obtained in the present study. With the long stimulus duration, the MMN peak latency and the RT to targets in the corresponding active discrimination task were longer for missing-fundamental tones than for pure tones by similar amounts of time. In contrast, in the present study, no correlation was found between the presence of missing-fundamental pitch information in the neuronal population generating the N1 potential and voluntary discrimination of these complex tones. This finding supports the view that the N1 generating mechanisms are not a part of the sequence of brain events directly underlying perception (Walter, 1964; Davis and Zerlin, 1966; Näätänen and Picton, 1987; Näätänen, 1992).

It is possible that the missing-fundamental pitch information reflected by the supratemporal N1 is established in subcortical parts of the afferent auditory system. This notion is supported by observations of neural activity related to the missing-fundamental pitch throughout the ascending auditory pathway (Smith *et al.*, 1978; Greenberg *et al.*, 1987; Galbraith, 1994). Furthermore, the subcortical pitch-related neural activity does not require long acoustic samples of missing-fundamental tones (e.g., a frequency-following response of the fundamental commences in the brainstem as soon as 15 ms after the onset of a missing-fundamental tone; Smith *et al.*, 1978; Greenberg *et al.*, 1987). At the cortical level, tonotopically organized neuronal populations have been shown to underlie the Pa component (Pantev *et al.*, 1995), one of the earliest exogenous (''obligatory,'' stimulus driven) auditory cortical evoked responses (peaking at about 30 ms from stimulus onset). In addition, the peak latency of this, as well as the other ''middle-latency'' auditory evoked potentials (Na, Nb, and Pb) was shorter in response to 4-kHz than to 250-Hz tone bursts (Woods *et al.*, 1995). This result provides further evidence linking the observed spectral and virtual pitch specific N1 latency effects to earlier phases of pitch processing. Thus there seems to exist an unbroken chain of stages along the auditory afferent pathway establishing the pitch of missing-fundamental tones with very short delay from the onset of stimulation and encoding it by the same tonotopical code as that for the spectral pitch. The relatively early onset and tonotopical nature of the exogenous supratemporal N1 component are compatible with the notion that the supratemporal N1 neurons could receive their input from this system. The present suggestion does not necessarily contradict Pantev *et al.*'s (1996) findings demonstrating the ''central'' origin of the missing-fundamental pitch code reflected by the N1m as the auditory pathways from the two ears converge already in the medial superior olive, the second ''relay station'' of ascending auditory information in the brainstem (see, e.g., Harrison, 1978).

The missing-fundamental pitch-extraction process subserving the auditory stimulus representations of the MMN system requires considerably longer (longer than 150 ms) acoustic samples than that needed for preattentively discriminating pure tones of even only slightly differing frequencies (30 ms in Paavilainen *et al.*, 1993). The relatively long duration of this process is also suggested by the long (ca. 250 ms) peak latency of the MMN to missing-fundamental pitch deviants (see also Winkler *et al.*, 1995). These features are compatible with the assumption that the missing-fundamental pitch is resolved by some central pattern-recognition process for subsequent storage in auditory sensory memory.

V. CONCLUSIONS

The problem of resolving the missing-fundamental pitch has important implications for the theory of pitch perception. The present study demonstrated that two separate codes of the missing-fundamental pitch exist in the auditory cortex. These codes are determined by different pitch-extraction processes. The auditory stimulus representations involved in preattentive change detection (reflected by the MMN component) seem to correspond to those subserving voluntary discrimination. Establishing this code takes considerable time and requires long samples of the acoustic input. The presence of missing-fundamental pitch information in the tonotopically organized neuronal population generating the supratemporal N1 did not correspond to the voluntary discrimination of missing-fundamental tones. The pitch information reflected by the supratemporal N1 is probably established by subcortical mechanisms in the early phase of auditory processing shortly after the stimulus onset. Therefore, the present results suggest the existence of redundant mechanisms in resolving the missing-fundamental pitch.

ACKNOWLEDGMENTS

This study was supported by the National Scientific Research Fund of Hungary (OTKA T006967), the Academy of Finland, and the Finnish Psychological Society. The authors thank Dr. István Czigler, Dr. Márk Molnár, Dr. Walter Ritter, and Dr. Hannu Tiitinen for their helpful comments and suggestions.

[1]This process might be feature integration, as, unlike the supratemporal N1, MMN has been shown to reflect the feature conjunctions present in a given stimulus (Gomes *et al.*, in press, see Ritter *et al.*, 1995), as well as the temporal structure of complex sounds (e.g., Winkler and Schröger, 1995).

Aaronson, D., and Walls, B. (**1987**). ''Extension of Grier's computational formula for A′ and B″ to below-chance performances,'' Psychol. Bull. **102,** 439–442.

Alho, K. (**1995**). ''Cerebral generators of mismatch negativity (MMN) and its magnetic counterpart (MMNm) elicited by sound changes,'' Ear Hear. **16**, 38–51.

Barry, R. J., Cocker, K. I., Anderson, J. W., Gordon, E., and Rennie, C. (**1992**). ''Does the N 100 evoked potential really habituate? Evidence from a paradigm appropriate to a clinical setting,'' Int. J. Psychophysiol. **13**, 9–16.

Butler, R. A. (**1972**). ''The auditory evoked response to stimuli producing periodicity pitch,'' Psychophysiology **9,** 233–237.

Carlyon, R. P., and Shackleton, T. H. (**1994**). ''Comparing the fundamental frequencies of resolved and unresolved harmonics: Evidence for two mechanisms?,'' J. Acoust. Soc. Am. **95,** 3541–3554.

Clarkson, M. G., and Clifton, R. K. (**1985**). ''Infant pitch perception: Evidence for responding to pitch categories and the missing fundamental,'' J. Acoust. Soc. Am. **77,** 1521–1528.

Cowan, N. (**1995**). *Attention and Memory. An Integrated Framework* (Oxford U.P., Oxford), pp. 70–75.
Davis, H., and Zerlin, S. (**1966**). ''Acoustic relations of the human vertex potential,'' J. Acoust. Soc. Am. **39,** 109–116.
de Boer, E. (**1976**). ''On the 'residue' and auditory pitch perception,'' in *Handbook of Sensory Physiology: Vol. 3*, edited by W. D. Keidel and W. D. Neff (Springer-Verlag, New York), pp. 479–583.
Dixon, W. J. (**1985**). *BMDP Statistical Software Manual* (University of California, Berkeley).
Elberling, C., Bak, C., Kofoed, B., Lebech, J., and Saermak, K. (**1982**). ''Auditory magnetic fields. Source localization and 'tonotopical organization' in the right hemisphere of the human brain,'' Scand. Audiol. **11,** 61–65.
Evans, E. F. (**1977**). ''Frequency selectivity at high signal levels of single units in cochlear nerve and nucleus,'' in *Hearing-Psychophysics and Physiology of Hearing*, edited by E. F. Evans and J. P. Wilson (Academic, London), pp. 185–196.
Galbraith, G. C. (**1994**). ''Two-channel brain-stem frequency-following response to pure tone and missing fundamental stimuli,'' Electroencephalogr. Clin. Neurophysiol. **92,** 321–330.
Goldstein, J. L. (**1973**). ''An optimum processor theory for the central formation of the pitch of complex tones,'' J. Acoust. Soc. Am. **54,** 1496–1516.
Gomes, H., Bernstein, R., Ritter, W., Vaughan, Jr., H. G., and Miller, J. (**in press**). ''Storage of feature conjunctions in transient auditory memory,'' Psychophysiology.
Greenberg, S., Marsh, J. T., Brown, W. S., and Smith, J. C. (**1987**). ''Neural temporal coding of low pitch. I. Human frequency-following responses to complex tones,'' Hearing Res. **25,** 91–114.
Grier, J. B. (**1971**). ''Nonparametric indexes for sensitivity and bias,'' Psychol. Bull. **75,** 424–429.
Halgren, E., Baudena, P., Clarke, J. M., Heit, G., Liégeois, C., Chauvel, P., and Musolino, A. (**1995**). ''Intracerebral potentials to rare target and distractor auditory and visual stimuli: I. superior temporal plane and parietal lobe,'' Electroencephalogr. Clin. Neurophysiol. **94,** 191–220.
Hari, R., Hämäläinen, M., Ilmoniemi, R., Kaukoranta, E., Reinikainen, K., Salminen, J., Alho, K., Näätänen, R., and Sams, M. (**1984**). ''Responses of the primary auditory cortex to pitch changes in a sequence of tone pips: Neuromagnetic recordings in man,'' Neurosci. Lett. **50,** 127–132.
Hari, R., Kaila, K., Katila, T., Tuomisto, T., and Varpula, T. (**1982**). ''Interstimulus interval dependence of the auditory vertex response and its magnetic counterpart: Implications for their neural generation,'' Electroencephalogr. Clin. Neurophysiol. **54**, 561–569.
Harrison, J. M. (**1978**). ''Functional properties of the auditory system of the brain stem,'' in *Handbook of Behavioral Neurobiology: Vol. 1, Sensory Integration*, edited by R. B. Masterton (Plenum, New York), pp. 409–458.
Horst, J. W., Javel, E., and Farley, G. R. (**1986**). ''Coding of spectral fine structure in the auditory nerve. I. Fourier analysis of period and interspike interval histograms,'' J. Acoust. Soc. Am. **79,** 398–416.
Houtsma, A. J. M., and Goldstein, J. L. (**1972**). ''The central origin of the pitch of complex tones: evidence from musical interval recognition,'' J. Acoust. Soc. Am. **51,** 520–529.
Houtsma, A. J. M., and Smurzinsky, J. (**1990**). ''Pitch identification and discrimination for complex tones with many harmonics,'' J. Acoust. Soc. Am. **87,** 304–310.
Jacobson, G. P., Lombardi, D. M., Gibbons, N. D., Ahmad, B. K., and Newman, C. W. (**1992**). ''The effects of stimulus frequency and recording site on the amplitude and latency of multichannel cortical auditory evoked potential (CAEP) component N1,'' Ear Hear. **13**, 300–306.
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
Krumhansel, C. L. (**1990**). *Cognitive Foundations of Music Pitch* (Oxford U.P., Oxford), pp. 112–119.
Lounasmaa, O. V., Hari, R., Joutsiniemi, S.-L., and Hämäläinen, M. (**1989**). ''Multi-SQUID recordings of human cerebral magnetic fields may give information about memory processes,'' Europhys. Lett. **9,** 603–608.
Lu, S. T., Hari, R., and Sams, M. (**1989**). ''Neuromagnetic responses to infrequent changes in periodicity pitch,'' in *Advances in Biomagnetism*, edited by S. J. Williamson, M. Hoke, G. Stroink, and M. Kotani (Plenum, New York), pp. 117–120.
Lyytinen, H., Blomberg, A. P., and Näätänen, R. (**1992**). ''Event-related potentials and autonomic responses to a change in unattended auditory stimuli,'' Psychophysiology **29,** 523–534.
Näätänen, R. (**1992**). *Attention and Brain Function* (Erlbaum, Hillsdale, NJ).
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 on evoked potential reinterpreted,'' Acta Psychol. **42,** 313–329.
Näätänen, R., Paavilainen, P., Alho, K., Reinikainen, K., and Sams, M. (**1989**). ''Do event-related potentials reveal the mechanism of the auditory sensory memory in the human brain?,'' Neurosci. Lett. **98,** 217–221.
Näätänen, R., and Picton, T. W. (**1987**). ''The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure,'' Psychophysiology **24,** 375–425.
Näätänen, R., Sams, M., Alho, K., Paavilainen, P., Reinikainen, K., and Sokolov, E. N. (**1988**). ''Frequency and location specificity of the human vertex N1 wave,'' Electroencephalogr. Clin. Neurophysiol. **69,** 523–531.
Novak, G. P., Ritter, W., and Vaughan, Jr., H. G. (**1992**). ''The chronometry of attention-modulated processing and automatic mismatch detection,'' Psychophysiology **29,** 412–430.
Paavilainen, P., Jiang, D., Lavikainen, J., and Näätänen, R. (**1993**). ''Stimulus duration and the sensory memory trace: An event-related potential study,'' Biol. Psychol. **35,** 139–152.
Pantev, C., Bertrand, O., Eulitz, C., Verkindt, C., Hampson, S., Schuierer, G., and Elbert, T. (**1995**). ''Specific tonotopic organizations of different areas of the human auditory cortex revealed by simultaneous magnetic and electric recordings,'' Electroencephalogr. Clin. Neurophysiol. **94,** 26–40.
Pantev, C., Elbert, T., Ross, B., Eulitz, C., and Terhardt, E. (**1996**). ''Binaural fusion and the representation of virtual pitch in the human auditory cortex,'' Hearing Res. **100,** 164–170.
Pantev, C., Hoke, M., Lehnertz, K., Lütkenhöner, B., Anogianakis, G., and Wittkowski, W. (**1988**). ''Tonotopic organization of the human auditory cortex revealed by transient auditory evoked magnetic fields,'' Electroencephalogr. Clin. Neurophysiol. **69,** 160–170.
Pantev, C., Hoke, M., Lütkenhöner, B., and Lehnertz, K. (**1989**). ''Tonotopic organization of the auditory cortex: Pitch versus frequency representation,'' Science **246,** 486–488.
Parasuraman, R., Richer, F., and Beatty, J. (**1982**). ''Detection and recognition: Concurrent processes in perception,'' Percept. Psychophys. **31,** 1–12.
Ragot, R., and Lepaul-Ercole, R. (**1996**). ''Brain potentials as objective indexes of auditory pitch extraction of harmonics,'' NeuroReport **7**, 905–909.
Ritter, W., Vaughan, Jr., H. G., and Costa, L. D. (**1968**). ''Orienting and habituation to auditory stimuli: A study of short-term changes in averaged evoked responses,'' Electroencephalogr. Clin. Neurophysiol. **25,** 550–556.
Ritter, W., Deacon, D., Gomes, H., Javitt, D. C., and Vaughan, H. G., Jr. (**1995**). ''The mismatch negativity of event-related potentials as a probe of transient auditory memory: A review,'' Ear Hear. **16**, 52–67.
Roberts, T. P. L., and Poeppel, D. (**1996**). ''Latency of auditory evoked M100 as a function of tone frequency,'' NeuroReport **7**, 1138–1140.
Romani, G. L., Williamson, S. J., and Kaufman, L. (**1982**). ''Tonotopic organization of the human auditory cortex,'' Science **216,** 1339–1340.
Sams, M., and Salmelin, R. (**1994**). ''Evidence of sharp tuning in the human auditory cortex,'' Hearing Res. **75,** 67–74.
Scharf, B., and Houtsma, A. J. M. (**1986**). ''Audition II: Loudness, pitch, localization, aural distortion, pathology,'' in *Handbook of Perception and Human Performance: Vol. I, Sensory Processes and Perception*, edited by K. R. Boff, L. Kaufman, and J. P. Thomas (Wiley, New York), pp. 15.2–60.
Scherg, M., Vajsar, J., and Picton, T. W. (**1989**). ''A source analysis of the late human auditory evoked potentials,'' J. Cogn. Neurosci. **1**, 336–355.
Schouten, J. F. (**1949**). ''The residue and the mechanism of hearing,'' Proc. Kon. Ned. Akad. Wet. **43,** 991–999.
Schouten, J. F. (**1970**). ''The residue revisited,'' in *Frequency Analysis and Periodicity Detection in Hearing*, edited by R. Plomp and G. F. Smoorenburg (Sijthoff, Lieden), pp. 41–54.
Schouten, J. F., Ritsma, R. J., and Cardoso, B. L. (**1962**). ''Pitch of the residue,'' J. Acoust. Soc. Am. **34,** 1418–1424.
Seebeck, A. (**1841**). ''Beoubachtungen über einige Bedingungen der Entstehun von Tönen,'' Ann. Phys. Chem. **53**, 417–436.
Smith, J. C., Marsh, J. T., Greenberg, S., and Brown, W. S. (**1978**). ''Human auditory frequency-following responses to a missing fundamental,'' Science **201,** 639–641.
Squires, K. C., Squires, N. K., and Hillyard, S. A. (**1975**). ''Decision-related cortical potentials during an auditory signal detection task with cued ob-

servation intervals,'' J. Exp. Psychol.: Human Perc. Perf. **1**, 268–279.
Terhardt, E. (**1974**). ''Pitch, consonance, and harmony,'' J. Acoust. Soc. Am. **55,** 1061–1069.
Tiitinen, H., May, P., Reinikainen, K., and Näätänen, R. (**1994**). ''Attentive novelty detection in humans is governed by pre-attentive sensory memory,'' Nature (London) **372,** 90–92.
Tomlinson, R. W., and Schwarz, D. W. (**1988**). ''Perception of the missing fundamental in nonhuman primates,'' J. Acoust. Soc. Am. **84,** 560–565.
Vaughan, Jr., H. G., and Ritter, W. (**1970**). ''The sources of auditory evoked responses recorded from the human scalp,'' Electroencephalogr. Clin. Neurophysiol. **28,** 360–367.
Walter, W. G. (**1964**). ''The convergence and interaction of visual, auditory, and tactile responses in human nonspecific cortex,'' Ann. (N.Y.) Acad. Sci. **112,** 320–361.
Wightman, F. L. (**1973**). ''The pattern-transformation model of pitch,'' J. Acoust. Soc. Am. **54,** 407–416.
Winkler, I., Reinikainen, K., and Näätänen, R. (**1993**). ''Event-related brain potentials reflect traces of the echoic memory in humans,'' Percept. Psychophys. **53,** 443–449.
Winkler, I., and Schröger, E. (**1995**). ''Neural representation for the temporal structure of sound patterns,'' NeuroReport **6**, 690–694.
Winkler, I., Tervaniemi, M., Huotilainen, M., Ilmoniemi, R., Ahonen, A., Salonen, O., Standertskjöld-Nordenstam, C.-G., and Näätänen, R. (**1995**). ''From objective to subjective: pitch representation in the human auditory cortex,'' NeuroReport **6**, 2317–2320.
Woods, D. L., Alain, C., Covarrubias, D., and Zaidel, O. (**1995**). ''Middle latency auditory evoked potentials to tones of different frequency,'' Hearing Res. **85,** 69–75.

Harmonic fusion and pitch shifts of mistuned partials[a)]

Alain de Cheveigné
Laboratoire de Linguistique Formelle, CNRS/Université Paris 7, 2 place Jussieu, case 7003, 75251, Paris, France and ATR Human Information Processing Research Laboratories, 2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-02, Japan

(Received 30 May 1996; accepted for publication 14 April 1997)

A model is proposed to explain pitch shifts of components mistuned from a harmonic series [W. M. Hartmann and S. L. Doty, ''On the pitches of the components of a complex tone,'' J. Acoust. Soc. Am. **99**, 567–578 (1996)]. An internal random variable depends on the frequency of the partial and determines both its pitch and the probability that it will fuse within the complex. The model accounts for the sign of the pitch shifts (positive for positive mistunings, negative for negative mistunings), the fact that they saturate and become smaller for large mistunings, and, given reasonable assumptions on the underlying distributions, their overall shape and magnitude. No assumptions are required concerning the pitch extraction mechanism (place, time, etc.) other than that a common source of variability affects cues to both pitch and harmonic fusion. © *1997 Acoustical Society of America.* [S0001-4966(97)07607-8]

PACS numbers: 43.66.Ba, 43.66.Hg [JWH]

INTRODUCTION

Major aspects of pitch perception are predicted equally well by different pitch perception models. For this reason, minor effects (such as pitch shifts) take on great theoretical importance, as they alone can reveal the presence of particular underlying mechanisms. Hartmann *et al.* (1986, 1990) asked subjects to match the pitch of a mistuned component of a harmonic complex to that of an isolated tone of similar frequency. The task was relatively easy for components of low rank and large mistunings, and more difficult for higher ranks and smaller mistunings. An interesting aspect of their data, elaborated by Hartmann and Doty (1996), was that mistunings were systematically overestimated by the subjects. The pitch of the mistuned component was shifted by up to several percent, in a direction that matched that of the mistuning. The shift tended to peak at about 4% mistuning and decrease beyond.

On the basis of these shifts, Hartmann and Doty (1996) argued against the model of Terhardt (1979) that predicts shifts in one direction whatever the mistuning. They proposed instead two versions of a time-domain model based on peaks of interspike interval (ISI) histograms. In the first version, the histograms were supposed to be gathered from fibers of the same characteristic frequency as the mistuned component. The predicted shifts did not match those observed. In a second version, the histograms were derived from a higher-frequency region (because of the upward spread of masking). The second version successfully accounted for major aspects of the shifts (sign, magnitude) of all components of rank greater than 1. However it could not account for shifts observed at the fundamental, nor could it account for the saturation and decrease of the pitch shift beyond 4% mistuning.

Another possible explanation is that the shifts are a side effect of harmonic fusion. Partials that match the harmonic series of a complex tone tend to fuse with it, whereas mistuned partials segregate and are easier to hear (Moore *et al.*, 1985, 1986; Hartmann, 1988; Hartmann *et al.*, 1986, 1990; Martens, 1981). When the partials of a concurrent vowel pair follow the same harmonic series, the pair is heard as s single source more often than when there is a ΔF_0, or when either vowel is inharmonic (de Cheveigné *et al.*, 1997). If, in a hypothetical variant of Hartmann and Doty's experiment, the partial's frequency had been randomly distributed, successful matches would have been less likely on trials for which the partial fell closer to the harmonic series, with the result that the distribution of pitch matches would be biased in a way similar to that observed. In Hartmann and Doty's experiment frequencies were fixed and not random, but the same reasoning may be applied to an internal representation used by both pitch and harmonic fusion and randomly distributed from trial to trial. This hypothesis is explored in the present paper.

I. MODEL

We assume that a sine-wave component gives rise to an internal correlate randomly distributed from trial to trial. This correlate determines both the probability of fusion, and the pitch perceived on trials for which the partial is not fused. It varies on a scale x about a mean x_0 that is proportional to the frequency f of the component, expressed as percentage mistuning from the harmonic series. ''Trial'' may be interpreted as meaning a pitch match, some of which were successful, others not (Hartmann and Doty, 1996). It might also mean an individual presentation of the stimulus, each of which gave rise to a slightly different pitch. Or else a ''trial'' might be an elementary frequency estimate that contributed to the pitch with a probability (or weight) depending on how close it fell to the harmonic series.

Let us denote $a(x)$ as the probability density of x, distributed around its mean x_0 [Fig. 1(A)], and $b(x)$ as the probability of harmonic fusion as a function of x [Fig. 1(B)].

[a)]Parts of this work were presented at the Spring 1997 meeting of the Acoustical Society of Japan.

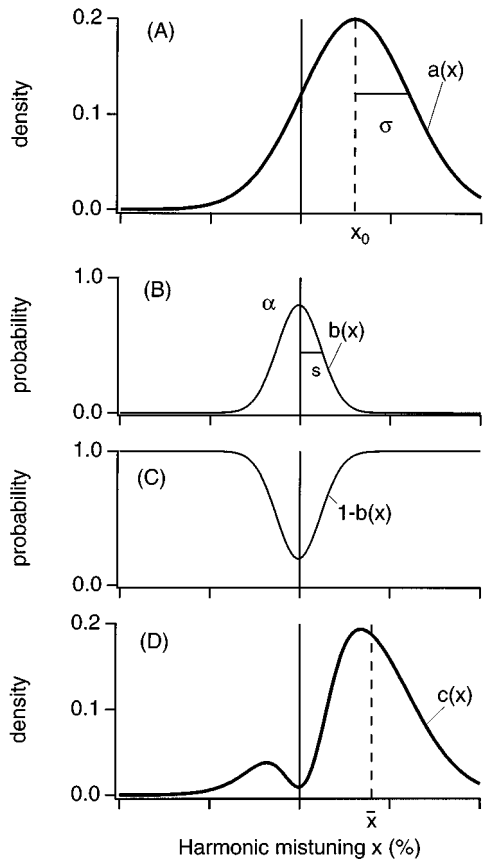


FIG. 1. (A) Distribution of x for all trials. The vertical dotted line marks the center of gravity of the distribution, x_0. (B) Probability of harmonic fusion. (C) Probability of a successful pitch match. (D) Distribution of x for successful matches. The vertical dotted line marks the center of gravity $\overline{x}$ of this distribution. The distance from the dotted line in (A) represents the pitch shift.

Because of fusion, pitch matches are less likely to be successful when x falls near zero [Fig. 1(C)]. The distribution of x for successful pitch matches [denoted as $c(x)$] is therefore distorted relative to the original distribution $a(x)$ [Fig. 1(D)]. Its center of gravity is shifted in the same direction as the mistuning, hence the pitch shift in that direction. With large mistunings the distribution of x is less affected by harmonic fusion, and pitch shifts should decrease. Both aspects were evident in the data of Hartmann and Doty (1996).

With a reasonable choice of the distribution of x and of the probability of fusion the shift may be calculated. Let us assume that x is normally distributed about its mean x_0:

$$a(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-(x-x_0)^2/2\sigma^2}. \tag{1}$$

We suppose also that the probability of fusion conditional on x is shaped like a Gaussian centered at zero mistuning:

$$b(x)=\alpha e^{-x^2/2s^2}. \tag{2}$$

The parameter s controls the ''width'' of this slot of the harmonic sieve (Duifhuis *et al.*, 1982) and α, with a value between 0 and 1, its ''permeability.'' Together with σ, the model thus has three parameters. The distribution of successful matches is

$$c(x)=Aa(x)(1-b(x)), \tag{3}$$

where the normalization factor A ensures that the distribution sums to 1. This may be rewritten as the weighted sum of two normal distributions:

$$c(x)=A\left(\frac{1}{\sigma\sqrt{2\pi}}e^{-e(x-x_0)^2/2\sigma^2}-B\frac{1}{D\sqrt{2\pi}}e^{-(x-C)^2/2D^2}\right)$$

with

$$B=\frac{\alpha s}{\sqrt{s^2+\sigma^2}}e^{-x_0^2/2(s^2+\sigma^2)},$$

$$C=\frac{fs^2}{s^2+\sigma^2},\quad D=\frac{s\sigma}{\sqrt{s^2+\sigma^2}}.$$

Noting that a normal distribution has an integral of 1, we have

$$A=1/(1-B).$$

The mean $\overline{x}$ of the distribution of x for successful matches is

$$\overline{x}=\int_{-\infty}^{\infty}xc(x)dx=A(x_0-BC). \tag{4}$$

The pitch shift is the difference between the mean $\overline{x}$ of the distorted distribution and the mean x_0 of the original distribution. This can be calculated as

$$\overline{x}-x_0=x_0\frac{\sigma^2}{s^2+\sigma^2}\Bigg/\left[\frac{\sqrt{s^2+\sigma^2}}{\alpha s}e^{x_0^2/2(s^2+\sigma^2)}-1\right]. \tag{5}$$

Figure 2 illustrates the range of shift patterns that are obtained with various values of the three parameters α (''permeability'' of the harmonic sieve), s (width of harmonic sieve), and σ (standard deviation of the internal variable). In Fig. 2(A) parameter α was fixed at 1, and parameters s and σ were varied together from 1% to 10%. All curves have the same slope at the origin: When $s=\sigma$ this slope depends only on α. The dependency on α is illustrated in Fig. 2(B) for $s=\sigma=3\%$. Finally, in Fig. 2(C), α was equal to 1, and the geometric mean of s and σ was kept constant and equal to 3%, while the ratio s/σ was varied from 0.3 ($s=1\%$, $\sigma=9\%$) to 3 ($s=9\%$, $\sigma=1\%$) in 10 logarithmically spaced steps. When s/σ is small, shifts tend to increase gradually with mistuning, whereas when it is large they peak at a small mistuning and then decrease beyond. The model can thus produce a wide range of shift patterns that are all symmetrical about the origin.

II. COMPARISON WITH EXPERIMENTAL DATA

Symbols in Fig. 3(A) represent shifts observed by Hartmann and Doty (1996) for the fifth harmonic at a low level (28 dB per harmonic). The line represents shifts produced by the model for $\alpha=0.8$ and $s=\sigma=3\%$. These parameter values were chosen to produce a good fit ''by eye.'' The shifts have the same sign as the mistuning and peak at about 4% mistuning and decrease thereafter, as observed experimentally.

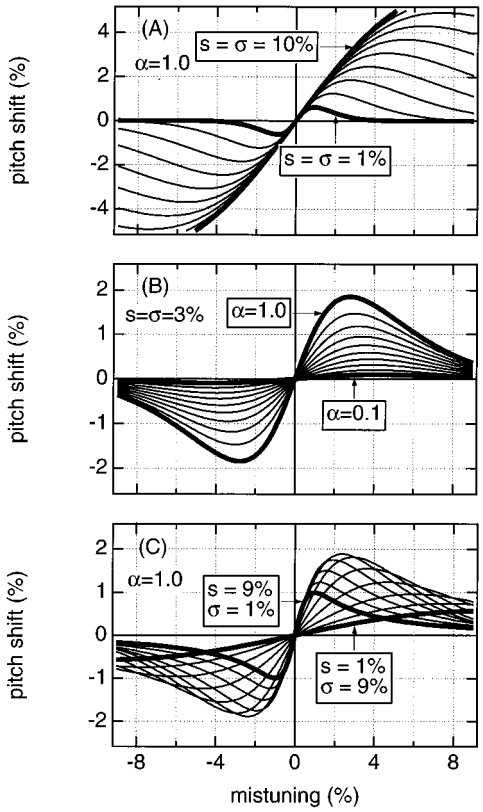


FIG. 2. (A) Pitch shifts produced by the model as a function of mistuning, for $\alpha=1.0$ and for values of $s=\sigma$ varying between 1% and 10%. (B) Same, for $s=\sigma=3\%$ and α varying between 0.1 and 1.0. (C) Same, for s/σ varying in logarithmically spaced steps from 0.3 to 3, while $\alpha=1.0$ and the geometric mean of s and σ remains constant and equal to 3%.

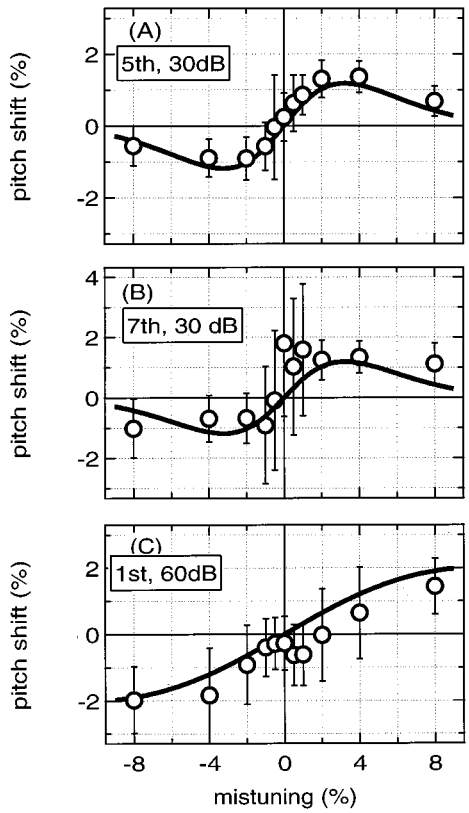


FIG. 3. (A) Symbols: pitch shifts observed for the fifth harmonic at a level of 28 dB per harmonic (Hartmann and Doty, 1996). Line: shifts produced by the model assuming $s=\sigma=3\%$ and $\alpha=0.8$. (B) As in (A), for the seventh harmonic at a level of 28 dB per harmonic. (C) Symbols: pitch shifts observed for the fundamental at a level of 58 dB per harmonic. Line: shifts produced by the model assuming $s=\sigma=8\%$ and $\alpha=0.6$.

Symbols in Fig. 3(B) represent shifts observed for the seventh harmonic at 28 dB per harmonic. The line represents shifts produced by the model using the same set of parameters as in Fig. 3(A). The ''fit'' is less good, but the model-produced shifts nevertheless fall within experimental error bars. The same can be said of most other conditions: harmonics 2, 3, and 4 at 28 dB per harmonic, and harmonics 3, 4, 5, and 11 at 58 dB per harmonic. Other conditions require a different choice of parameters. The somewhat larger shifts observed by Hartmann and Doty at harmonics 9 and 11 at 28 dB per harmonic, and harmonic 7 at 58 dB per harmonic (not shown) can be accommodated by assuming $s=\sigma=4\%$, and $\alpha=0.8$ as before. The monotonically increasing shifts at the fundamental (58 dB per harmonic) can be accounted for by assuming $s=\sigma=8\%$ and $\alpha=0.6$ [Fig. 3(C)]. Finally, shifts at the second harmonic at 58 dB per harmonic (not shown) can be modeled by assuming $s=\sigma=6\%$ and $\alpha=0.6$. Thus to account for the entire set of Hartmann and Doty's data one must assume that the parameters vary somewhat between harmonics and level, which is perhaps not surprising. Given experimental error and the tradeoffs between parameters, it is however difficult to make precise estimates of each parameter.

In all three examples shown in Fig. 3 there was a tendency for experimental data points to mostly fall above or below the curve predicted by the model. Systematic shifts of up to 2% were observed in a mistuned harmonic experiment by Lin and Hartmann (1996), who found considerable variation between conditions and subjects. The model makes predictions that are symmetric relative to the origin, and thus we cannot explain such systematic shifts. They might be explained by a response bias such as the tendency to make pitch adjustments slightly sharp as observed by Jarosewski (1992), or by shifts caused by neighboring partials as predicted by Terhardt (1979). Lin and Hartmann argue against this last interpretation, as the shifts do not seem to require the presence of either neighboring partial. A ''sideways'' discrepancy might be explained within the context of our model by assuming that the harmonic sieve is slightly mistuned. Apart from such systematic shifts, the model accounts for major aspects of Hartmann and Doty's data. Pitch shifts have the same sign as the mistuning, and with a reasonable choice of parameters they peak at about 4% mistuning and decrease beyond. This last aspect, highly reproducible according to Hartmann and Doty, was not explained by their model. Shifts at the fundamental, also not accounted for by their model, are explained here in the same fashion as for other harmonics.

III. SHIFTS OF THE LOW PITCH OF THE COMPLEX

Darwin *et al.* (1994) presented listeners with a complex tone that had a mistuned fourth harmonic. Instead of matching the pitch of the mistuned partial, as in Hartmann and

Doty's experiments, the subjects matched the low pitch of the entire complex. That pitch varied with the frequency of the partial up to about 3% mistuning, at which point it had shifted by about 0.5%. With further mistuning the low pitch shifted back, and by about 8% mistuning it had resumed its initial value, in agreement with the original observations of Moore *et al.* (1985). The pitch could be accurately modeled as

$$F_0 = a + k\Delta f e^{-\Delta f^2/2s^2}, \tag{6}$$

where F_0 is the frequency of the tone that matches the low pitch, Δf is the mistuning, and a is a constant close to the nominal pitch of the complex. The constant k is the rate of change in low pitch relative to changes in frequency of the partial, and s determines the range over which mistunings affect the low pitch. For mistunings expressed in percentage, their values were about 0.2% and 3%, respectively. When plotted, the complex pitch shift (Darwin *et al.*, 1994, Fig. 1) looks very similar to the partial pitch shift [Fig. 3(A)]. One must keep in mind their different nature: a shift of the low pitch of the complex relative to the fundamental of the fixed harmonic series in the former case, a shift of the pitch of the partial relative to its *own* mistuned frequency in the later. The low pitch is as it were ''pulled'' by the mistuned partial, whereas the pitch of the partial is ''pushed'' away from the harmonic series.

It is tempting to view the pitch shift of the complex as the outcome of the same probabilistic process that produces pitch shifts of the partials. According to our model (Sec. I) the partial's pitch was available on trials for which there was *not* fusion. Suppose we assume that on trials for which there *was* fusion, the partial contributed to the low pitch instead. In that case, the decrease in complex pitch shift for larger mistunings [Eq. (6)] would simply reflect the fact that fusion becomes less probable, as described by Eq. (2) and illustrated in Fig. 1(B). The parameter s would be the same in both equations, and this would explain why the value found by Darwin *et al.* (1994) (3%) works well in our model. This hypothesis is attractive because it allows the principle of disjoint allocation (Bregman, 1990) to apply at the microscopic level of the ''trial,'' even if it seems to fail at the macroscopic level of pitch matches (Moore, 1987; Darwin and Carlyon, 1995). However, we know of no direct experimental evidence in favor of this hypothesis. As it is unnecessarily strong in the context of our model, we shall explore it no further.

Our model assumes that pitch shifts are the consequence of harmonic fusion rather than interactions between neighboring partials. In support of this assumption, Lin and Hartmann (1996) found shifts at harmonic positions within a harmonic complex, even when neighboring components were missing. For example, the pitch of a 200 Hz $\pm$8% partial showed shifts in the presence of a harmonic template formed by harmonics between 1000 and 3200 Hz. Similarly, pitch shifts at the third harmonic were significant in presence of a harmonic complex that lacked either the second or the fourth harmonic, or both. In general, however, the shifts tended to be larger if all background harmonics were present, in particular the harmonic of next-lowest rank relative to the mistuned partial.

Similar shifts have been observed for the pitch of tones preceded by a tone of similar frequency (Hartmann and Kanistanaux, 1979). Those authors proposed a model in which the excitation pattern evoked by a tone was distorted by adaptation, thus causing a pitch shift. If the quantity $a(x)$ in Eq. (3) is viewed as a deterministic excitation pattern, and $[1-b(x)]$ as distortion caused by the ''harmonic sieve,'' our model can be interpreted as formally similar to theirs, the effects of adaptation being replaced by those of the harmonic sieve. Kashino and Nishida (1996, 1997) found similarly shaped shifts in the localization of sources preceded by a source of similar ITD (interaural time delay). They interpreted them as resulting from a recalibration of the ITD-extraction mechanism, to enhance the representation of ITDs near that of the adaptor. This rescaling of perceptual dimensions would play a functional role similar to that proposed for visual aftereffects by Barlow (1990).

IV. CONCLUSION

Pitch shifts of mistuned partials were explained by postulating a noisy internal variable that determined both the pitch of the partial and the probability of fusion within the harmonic complex. Aspects accounted for were: (a) the sign of the pitch shifts (same as the mistuning); (b) their magnitude; (c) the fact that they peak at about 4% and decrease beyond; and (d) the fact that they affect individually audible harmonics of all ranks, including the fundamental. The explanation is congruent with the notion of harmonic fusion and segregation observed in many situations. No assumption was required concerning a particular model of pitch perception (place, time, etc.). This weakens the argument made by Hartmann and Doty (1996) in favor of ISI-based rather than other temporal or place mechanisms, on the grounds that they allow the shifts to be explained. Instead, a different but interesting conclusion may be drawn: Pitch and harmonic fusion must share a common physiological substrate subject to random variations from trial to trial. The alternative hypothesis, that pitch and fusion depend on separate mechanisms, each with its own source of variability, is contradicted by Hartmann and Doty's data.

ACKNOWLEDGMENTS

This work was carried out within a research agreement between ATR and CNRS. Thanks to Bill Hartmann for providing access to unpublished data and manuscripts, and for detailed comments on several versions of this manuscript, that also benefited from comments of one anonymous reviewer.

Barlow, H. B. (**1990**). ''A theory about the functional role and synaptic mechanism of visual after-effects,'' in *Vision: Coding and Efficiency*, edited by C. Blakemore (Cambridge U. P., Cambridge, England), pp. 363–375.

Bregman, A. S. (**1990**). *Auditory scene analysis* (MIT, Cambridge, MA).

Darwin, C. J., and Carlyon, R. P. (**1985**). ''Auditory grouping,'' in *Handbook of Perception and Cognition: Hearing*, edited by B. C. J. Moore (Academic, New York), pp. 387–424.

Darwin, C. J., Ciocca, V., and Sandell, G. J. (**1994**). ''Effects of frequency and amplitude modulation on the pitch of a complex tone with a mistuned harmonic,'' J. Acoust. Soc. Am. **95,** 2631–2636.

de Cheveigné, A., McAdams, S., and Marin, C. (**1997**). ''Concurrent vowel segregation. II: Effects of phase, harmonicity, and task,'' J. Acoust. Soc. Am. **101,** 2848–2856.

Duifhuis, H., Willems, L. F., and Sluyter, R. J. (**1982**). ''Measurement of pitch in speech: An implementation of Goldstein's theory of pitch perception,'' J. Acoust. Soc. Am. **71,** 1568–1580.

Hartmann, W. M., and Kanistanaux, D. C. (**1979**). ''The effect of a prior tone on the pitch of a short tone,'' in *Proceedings of the Research Symposium on the Psychology and Acoustics of Music,* edited by W. V. May (University of Kansas, Lawrence, KS), pp. 199–215.

Hartmann, W. M. (**1988**). ''Pitch perception and the segregation and integration of auditory entities,'' in *Auditory Function—Neurological Bases of Hearing,* edited by G. M. Edelman, W. E. Gall, and W. M. Cowan (Wiley, New York), pp. 623–645.

Hartmann, W. M., McAdams, S., and Smith, B. K. (**1986**). ''Matching the pitch of a mistuned harmonic in a complex tone,'' IRCAM annual report.

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.

Hartmann, W. M., and Doty, S. L. (**1996**). ''On the pitches of the components of a complex tone,'' J. Acoust. Soc. Am. **99,** 567–578.

Jarosewski, A. (**1992**). ''A study of constant (systematic) errors in pitch discrimination of short tone pulses,'' Acustica **77,** 106–110.

Kashino, M., and Nishida, S. (**1996**). ''Auditory localization aftereffects,'' J. Acoust. Soc. Am. **99,** 2596.

Kashino, M., and Nishida, S. (**1997**). ''Adaptation in the processing of interaural time differences revealed by the auditory localization aftereffects,'' J. Acoust. Soc. Am. (submitted).

Lin, J. L., and Hartmann, W. M. (**1996**). ''The pitch of mistuned harmonics: evidence for a template model,'' J. Acoust. Soc. Am. (submitted).

Martens, J. P. (**1981**). ''Comment on 'Algorithm for extraction of pitch and pitch salience from complex tonal signals' [J. Acoust. Soc. Am. **71**, 679–688 (1982)],'' J. Acoust. Soc. Am. **75,** 626–628.

Moore, B. C. J., Peters, R. W., and Glasberg, B. R. (**1985**). ''Thresholds for the detection of inharmonicity in complex tones,'' J. Acoust. Soc. Am. **77,** 1861–1867.

Moore, B. C. J., Peters, R. W., and Glasberg, B. R. (**1986**). ''Thresholds for hearing mistuned partials as separate tones in harmonic complexes,'' J. Acoust. Soc. Am. **80,** 479–483.

Moore, B. C. J. (**1987**). ''The perception of inharmonic complex tones,'' in *Auditory Processing of Complex Sounds,* edited by W. A. Yost and C. S. Watson (Erlbaum, Hillsdale, NJ), pp. 180–189.

Terhardt, E. (**1979**). ''Calculating virtual pitch,'' Hearing Res. **1,** 155–182.

Multiband detection of energy fluctuations[a)]

John H. Grose and Joseph W. Hall III
The University of North Carolina at Chapel Hill, Division of Otolaryngology/Head & Neck Surgery, 610 Burnett-Womack Building, Chapel Hill, North Carolina 27599-7070

(Received 3 March 1997; revised 14 April 1997; accepted 15 April 1997)

This study sought to characterize the integration of synchronous energy fluctuations across relatively independent spectral regions. The detection of four classes of signal was examined where each class of signal was associated with a change in energy over time. The four signal types were: (1) multicomponent tonal complexes, with each component centered in a narrow band of noise; (2) intensity increments in multiple narrow bands of noise; (3) intensity decrements in multiple narrow bands of noise; and (4) temporal gaps in multiple narrow bands of noise. Each signal type was examined in a separate experiment, although stimulus characteristics such as presentation level and frequency location were held constant. Experiment 1 confirmed that the detection of multitonal complexes masked by narrow bands of noise is linearly related to $\sqrt{N}$, where N is the number of signal components. Experiment 2 extended this to show that, when the signal was an increment in the level of a noise band carrier, threshold continued to be a linear function of $\sqrt{N}$, although threshold was about 2 dB higher than for a comparable tonal signal. Experiment 3 indicated that the detection of a decrement in energy in one or more noise bands was relatively poorer in terms of absolute level changes than was the detection of an energy increment in the same stimulus. Examination of psychometric functions for decrement detection suggested that performance improved by less than $\sqrt{N}$. Experiment 4 found that gap detection improved with increasing N but by a factor greater than the $\sqrt{N}$ expected on statistical grounds. Examination of the underlying psychometric functions confirmed this effect. The results of these experiments suggest that, for energy increments, the auditory system integrates information from across the spectrum in a statistically independent manner, at least over the frequency range examined here. This does not appear to be the case for the detection of energy gaps or decrements. © *1997 Acoustical Society of America.* [S0001-4966(97)02108-5]

PACS numbers: 43.66.Dc, 43.66.Fe, 43.66.Mk [WJ]

INTRODUCTION

The characterization of the auditory periphery as a bank of overlapping bandpass filters is an established tenet of psychoacoustic theory (Fletcher, 1940). However, the modeling of certain psychoacoustic functions does not appear to require the overt inclusion of peripheral filtering, although its existence may be tacitly acknowledged. One such area of psychoacoustic function is the temporal processing of broadband signals (for review, see Eddins and Green, 1995). For example, in the model accounting for temporal modulation transfer functions described by Viemeister (1979) and the variation of this model used by Forrest and Green (1987) to relate gap detection and modulation detection, the initial stage of the model is a bandpass filter whose bandwidth of 2000 or 4000 Hz clearly exceeds that of any single critical band (see also, Bacon and Viemeister, 1985; Green and Forrest, 1988). The tacit assumption of such models is that the features of the broadband signal which are important to the temporal process are, in some way, reconstituted following peripheral filtering; hence, the inclusion of critical band filtering in the model is not necessary. Even similar models which specifically include peripheral filtering, such as the models of temporal processing put forward by Buus and Florentine (1985) and Formby and Muir (1988), acknowledge the necessity of some sort of merging, or integration, of critical band output. The lack of a marked influence of peripheral auditory filtering on temporal processing has been a specific focus of attention (e.g., De Boer, 1986). However, the issue of interest in the present study is the characterization of the spectral integration. It was the purpose of this investigation to examine how information about energy changes is combined across frequency.

A number of paradigms are relevant to this issue. For example, the results of band-widening studies—such as the detection of temporal gaps as a function of stimulus bandwidth—provide one characterization of spectral integration (Buus and Florentine, 1985; Fitzgibbons, 1983; Florentine and Buus, 1983; Formby and Muir, 1988; Grose *et al.*, 1989; Shailer and Moore, 1983, 1985). However, a limitation of such paradigms is that it is difficult to distinguish between effects that occur at a subcritical band level and those which occur across critical bands. This investigation sought to examine the spectral integration of complex signals which were distributed across frequency in such a manner that each constituent of the signal was relatively isolated in frequency, and hence processed in relatively independent channels.

The integration of information across frequency was examined using four tasks involving multiple narrow bands of

[a)]Portions of this paper were presented at the 129th Meeting of the Acoustical Society of America, 30 May–3 June 1995, Washington, DC [J. Acoust. Soc. Am. **97**, 3274(A) (1995)].

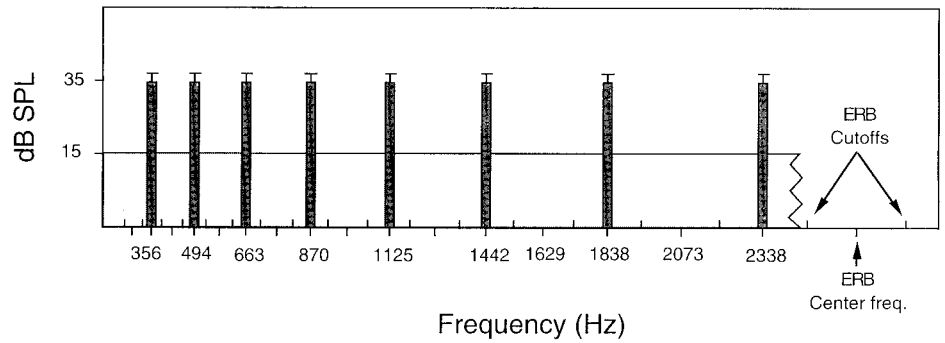


FIG. 1. Schematic spectra of eight-tone complex signal masked by its complementary eight-band masker complex. A continuous low-level background noise is also present. Abscissa gives center frequencies of successive ERBs (downward ticks), as well as upper and lower cutoffs of successive ERBs (upward ticks).

noise: (1) the detection of multicomponent tonal complexes masked by these bands; (2) the detection of energy increments in these bands; (3) the detection of energy decrements in these bands; and (4) the detection of temporal gaps in these bands. Each type of signal was examined in a separate experiment, although most stimulus characteristics, in particular presentation level and frequency location, were held constant across the experiments.

I. EXPERIMENT 1. TONE DETECTION AS A FUNCTION OF THE NUMBER OF TONES AND MASKING BANDS

The detection of a multicomponent tonal signal in noise as a function of the number of equally detectable components within the signal has been well-documented. For example, Buus *et al.* (1986) observed that the improvement in signal threshold as a function of component number could be expressed as

$$L_N = L_1 - 10\log(\sqrt{N}), \tag{1}$$

where L_N is threshold for an N-component signal, expressed in dB/component, and L_1 is threshold for a single-component signal. The linear improvement in threshold with $\sqrt{N}$ is consistent with the notion that sensitivity, d', is linearly related to signal level expressed in dB and that each component is treated as an independent source of information (Green, 1958). The paradigm used by Buus *et al.* (1986), as well as that in earlier studies (Green, 1958), involved measuring signal detection in a broadband masking noise. The purpose of this first experiment was to repeat the general approach of these studies but to use narrow bands of noise as the masker. Moreover, the number of masking bands of noise corresponded directly to the number of signal components such that a particular noiseband was present only if a signal component was also present at that center frequency. The purpose of this modification was to allow comparisons with the signal types used in the subsequent experiments. Because random noise remote in frequency from the signal is not expected to influence threshold (e.g., Fletcher, 1940), it was anticipated that the pattern of results from this experiment would conform to that found in earlier studies which used broadband maskers.

A. Method

1. Subjects

Eight normal-hearing subjects, one male and seven female, participated in the experiment. They ranged in age from 20 to 50 years. All subjects had pure-tone thresholds of 20 dB HL or better at octave frequencies from 250 to 8000 Hz. Each subject received at least 2 h of training prior to the collection of data until performance was stable; that is, until no systematic improvement in threshold was obvious.

2. Stimuli

Four signals were generated: (1) a 1125-Hz pure tone, (2) a two-tone complex consisting of 870 and 1125 Hz, (3) a four-tone complex consisting of 663, 870, 1125, and 1442 Hz, and (4) an eight-tone complex consisting of 356, 494, 663, 870, 1125, 1442, 1838, and 2338 Hz. Four corresponding maskers were also generated, each comprised of one or more 20-Hz-wide narrow bands of noise centered at the frequency of the respective signal component. As depicted in Fig. 1, the eight frequencies employed correspond to the center frequencies of auditory filters which, when expressed in equivalent rectangular bandwidths (ERBs) (Moore and Glasberg, 1983), are spectrally separated by one nonoverlapping ERB.

All stimuli were digitally generated at a sampling rate of 11 025 Hz and output through digital-to-analog converters (DACs) with 20-bit resolution (Digidesign Audiomedia and SoundAccelerator). Each noise band had components spaced at 0.1 Hz, yielding an overall periodicity of the noise of 10 s. Following output, the stimuli were antialias filtered at 4000 Hz (Kemo VBF, 90 dB/oct). The maskers were presented at a spectrum level of 35 dB/Hz (about 48 dB/noiseband). A continuous broadband background noise (0–5000 Hz) from a variable bandwidth noise generator (Hi-Mem Electronics IHR-QNG) was also presented at a spectrum level of 15 dB/Hz. All stimuli were delivered monaurally to the left ear via a Sony MDR-V6 headphone. Testing was conducted in a double-walled sound-attenuating chamber. Stimulus timing and response collection were controlled by a microcomputer.

3. Procedure

For each listener, threshold for each of the four signals was measured in its complementary masker. For example,

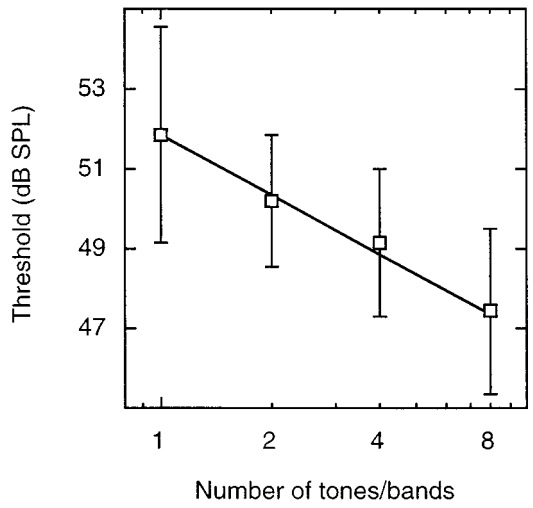


FIG. 2. Group mean results depicting signal threshold (dB SPL/component) as a function of the number of signal components. Error bars indicate ±1 standard deviation. See text for explanation of solid line.

threshold for the single-tone signal (1125 Hz) was measured in the presence of the single masking noise band centered at 1125 Hz, threshold for the two-tone signal (870+1125 Hz) was measured in the presence of the two noisebands centered at 870 and 1125 Hz, etc. For the multicomponent signals, the levels of the individual components were equal. A three-alternative, forced-choice (3AFC) procedure was used to measure threshold, incorporating a three-down, one-up stepping rule which converged upon the 79.4% correct level. In each of three intervals, the masker was gated on for 400 ms, including a 50-ms, cosine-squared rise–fall (Wilsonics BSIT). In one of the intervals, chosen at random, the signal was gated on and off synchronously with the masker. Subjects indicated their responses by means of a microterminal and feedback was provided after each trial. The step size for signal attenuation was initially set at 8 dB and this was reduced to 4 dB after two reversals in level direction; the step size was reduced to 2 dB after another two reversals, and the run was terminated after 12 reversals. The average of the final eight reversal levels was taken as the estimate of signal threshold for that run. For each listener, at least six estimates of threshold were obtained for each condition, and the final threshold value was taken as the mean of those six or more threshold estimates. The order of conditions was random across subjects.

B. Results and discussion

The group mean results were representative of the performance of the eight listeners and are displayed in Fig. 2. Mean signal threshold in dB/component is plotted against the number of signal components. A repeated measures analysis of variance (ANOVA) indicated a significant improvement in threshold with increasing number of signal components [$F(3,21)=30.03$; $p<0.01$]. *Post-hoc* contrasts indicated that the improvement in threshold with each doubling of component number was significant. The solid line in Fig. 2 is the function found by Buus *et al.* (1986) for tonal complex signals in a broadband noise masker [Eq. (1)]. It is clear that the observed data correspond very well to the prediction. This is noteworthy because, as mentioned earlier, the masking paradigm used in the Buus *et al.* (1986) study employed a broadband noise which was the same for all signals. In the present paradigm, the cumulative bandwidth of the masking noise was a function of the number of components present in any particular signal. Nevertheless, the expected improvement in signal threshold with increasing component number appears to hold even when the masker bandwidth is restricted to spectral regions directly centered on the signal components. The findings of this experiment therefore corroborate established results concerning the integration of signal energy from across the spectrum, but extend them to apply to the specific configuration wherein masker energy is present only in frequency regions of signal energy.

II. EXPERIMENT 2. INCREMENT DETECTION AS A FUNCTION OF THE NUMBER OF STANDARD NOISEBANDS

The detection of pure tones in noise can be viewed as a special case of intensity discrimination where the listener is detecting a change in intensity brought about by the addition of a tonal signal to a noise standard. This viewpoint motivated the study of Bos and de Boer (1966) which compared the detection of tonal signals in noise backgrounds of various bandwidths with intensity discrimination measures for those same noise bands. The intensity discrimination measures can be thought of as the detection of noise signals added to noise backgrounds. For relatively narrow bandwidths (less than about 160 Hz), performance for noise signals was 2 to 3 dB poorer than for tonal signals, in terms of $10\log(\Delta I/I)$. Bos and de Boer (1966) attributed this to the perceptually marked amplitude fluctuations which are inherent to narrow bands of noise and which made the detection of the noise signals more difficult than the tonal signals. Richards and Nekrich (1993) have also noted that the detection of tones in narrow bands of noise may involve strategies that are not applicable to the detection of simple level changes in those same noise bands. The purpose of experiment 2 was to measure intensity discrimination for the same noise bands which formed the maskers in experiment 1. The intent was to determine whether intensity discrimination changed as a function of the number of noise bands in a pattern similar to that found for multicomponent signal detection.

A. Method

1. Subjects

Eight normal-hearing subjects, one male and seven female, participated in experiment 2. None of the eight subjects had participated in experiment 1. The listeners in this experiment ranged in age from 22 to 33 years. All had pure-tone sensitivities of 20 dB HL or better at octave frequencies from 250 to 8000 Hz. Each subject received at least 2 h of training prior to the collection of data until performance was stable; that is, until no systematic improvement in threshold was obvious.

2. Stimuli

The stimuli for experiment 2 were the same eight narrow bands of noise that were used as maskers in experiment 1 (see Fig. 1). The level of the noise bands remained at 35 dB/Hz and a broadband background noise was again continuously present at a spectrum level of 15 dB/Hz.

3. Procedure

For each listener, increment detection threshold (intensity discrimination) was measured for four stimuli made up of some combination of the 20-Hz-wide narrow bands of noise. The four stimuli were the one-band, two-band, four-band and eight-band maskers of experiment 1. A 3AFC task was employed wherein an intensity increment was added to the standard stimulus in one of the three observation intervals. In each observation interval, the standard was gated on for 400 ms, including a 50-ms rise–fall time. In the target interval, the increment, which was 200 ms in duration, was added 100 ms after the onset of the observation interval. The increment resulted from an in-phase addition of a replica of the stimulus waveform to the standard. The amplitude of the replica waveform was the dependent variable and its addition to the standard waveform was controlled by an electronic gate (Wilsonics BSIT) having a cosine-squared rise–fall time 50-ms in duration. Implementation of the adaptive procedure and the number of replications per condition for each listener were the same as in experiment 1. Threshold was expressed in terms of 10 log $\Delta I/I$ where ΔI was derived as the intensity of an *independent* noise band which, when added to the standard, I, would have resulted in the same change in level (cf. Bos and de Boer, 1966).

B. Results and discussion

The group mean results were representative of the performance of the eight listeners and are displayed as diamonds in Fig. 3, which plots increment thresholds (10 log $\Delta I/I$) as a function of the number of bands. It is apparent that intensity discrimination improves as the number of noise bands increases, a result confirmed by a repeated measures ANOVA, which showed that the increment thresholds decreased significantly with band number [$F(3,21) = 41.56$; $p < 0.01$]. Moreover, *post-hoc* analysis indicated that the improvement was significant for each doubling of noise band number. The squares in Fig. 3 are the data from experiment 1 where the intensity of a tone at threshold has been related to the overall intensity of the noise band masking the signal (48 dB SPL). In agreement with Bos and de Boer (1966), it is apparent that the discrimination of tones in noise is superior to the discrimination of increments in the fluctuating noise bands themselves by about 2 dB. An ANOVA showed this difference to be significant [$F(1,14) = 5.42$; $p = 0.035$] and the lack of a significant interaction indicated that the difference remained constant for each noise band number. The fact that the signal duration was 400 ms in experiment 1 whereas the increment duration was 200 ms in experiment 2 is unlikely to impact greatly on this pattern since temporal integration for signal detection has largely reached an asymptote by a 200-ms duration (Green *et al.*, 1957; Watson and Gengel, 1969). The parallel improvement

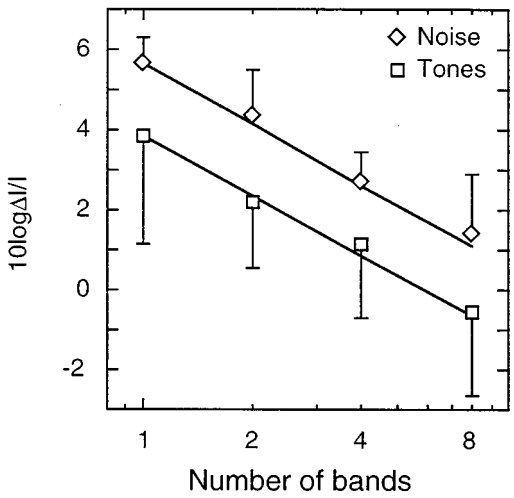


FIG. 3. Group mean results depicting intensity discrimination (10 log$\Delta I/I$) as a function of the number of signal components. Diamonds indicate intensity discrimination data for noise bands, with +1 standard deviation error bars; squares indicate data for tonal signals from experiment 1, with −1 standard deviation error bars. See text for explanation of solid lines.

in discrimination with increasing number of bands is notable and is highlighted by the solid diagonal lines which indicate the function

$$D_N = D_1 - 10 \log \sqrt{N}, \tag{2}$$

where D is the discrimination threshold expressed in terms of 10 log $\Delta I/I$ for N noise bands. The close fit of the data to this function strongly suggests that intensity discrimination is a linear function of $\sqrt{N}$, at least over the range of frequencies examined here. Thus each constituent is treated as an independent source of information.

III. EXPERIMENT 3. DECREMENT DETECTION AS A FUNCTION OF THE NUMBER OF STANDARD NOISEBANDS

Having established that intensity discrimination, or the detection of intensity increments, improves as a function of $\sqrt{N}$, it is of interest to consider the complementary situation, viz. the detection of decrements in intensity as a function of the number of stimulus constituents. A number of studies have explored the general relation between increment detection and decrement detection (de Boer, 1986; Forrest and Green, 1987; Macmillan, 1971; Moore and Peters, 1997; Moore *et al.*, 1993; Peters *et al.*, 1995). When expressed in terms of level change in dB (ΔL), Moore and Peters (1997) found that the absolute ΔL required for detecting a decrement in a mid- or high-frequency pure tone was usually greater than that required for detecting an increment. Hall and Grose (1991) also found that the ΔL for increments in narrow bands of noise was smaller than the ΔL for decrements. In contrast, de Boer (1986) found that thresholds for increment detection and decrement detection for pure-tone stimuli were the same as long as the duration of the energy change was greater than about 20 ms. Forrest and Green (1987) also found that the detectability of increments and

decrements in broadband noise was similar when expressed in terms of ΔL. However, they stressed that the interpretation of such results is strongly dependent on the choice of response metric; if their metric was instead the ratio of the change in amplitude relative to the standard amplitude ($\Delta A/A$), then for durations greater than about 10 ms, thresholds for decrements were lower than for increments. While the intent of these previous studies of increment and decrement detection has been primarily to model the sensitivity of the auditory system to intensity changes (positive and negative) without regard to spectral distribution of energy, the specific purpose of this experiment was to compare the spectral integration of intensity decrements to that for intensity increments.

A. Methods

1. Subjects

The same eight subjects that participated in experiment 2 participated in experiment 3.

2. Stimuli

The stimuli for experiment 3 were the same eight narrow bands of noise that were used as maskers in experiment 1 and as standards in experiment 2 (see Fig. 1). The level of the noise bands remained at 35 dB/Hz and the level of the broadband background noise remained at 15 dB/Hz.

3. Procedure

For each listener, decrement detection threshold was measured for the same four standard stimuli as in experiment 2. The procedure was in all respects the same as in experiment 2 except that a decrement, rather than an increment, was added to the standard in the signal interval. The decrement resulted from an antiphasic addition of a replica of the stimulus waveform to the standard. Again, threshold was expressed in terms of a derived ΔI which corresponded to the intensity of an *independent* noise band which, when subtracted from the standard, would have results in the same change in level.

B. Results and discussion

The group mean results were representative of the performance of the eight listeners and are displayed as squares in Fig. 4. The upper panel plots decrement detection in terms of 10 log $\Delta I/I$ as a function of the number of bands (squares). Here, all points fall below zero since the level of the inverted stimulus waveform which results in the decrement can never exceed the level of the standard (uninverted) waveform. The upper panel also replots the increment detection data of Fig. 3 for comparison (circles). The lower panel plots the data in terms of level change, or ΔL. Here ΔL for decrements corresponds to $10\log(1-\Delta I/I)$ and ΔL for increments corresponds to $10\log(1+\Delta I/I)$.

As mentioned earlier, measurements which deal with sensitivity to level changes in a sound can be quantified using a variety of interrelated metrics (cf. Green, 1988). Buus and Florentine (1991) argued that, at least for increment detection, ΔL is the most appropriate metric since d' is nearly

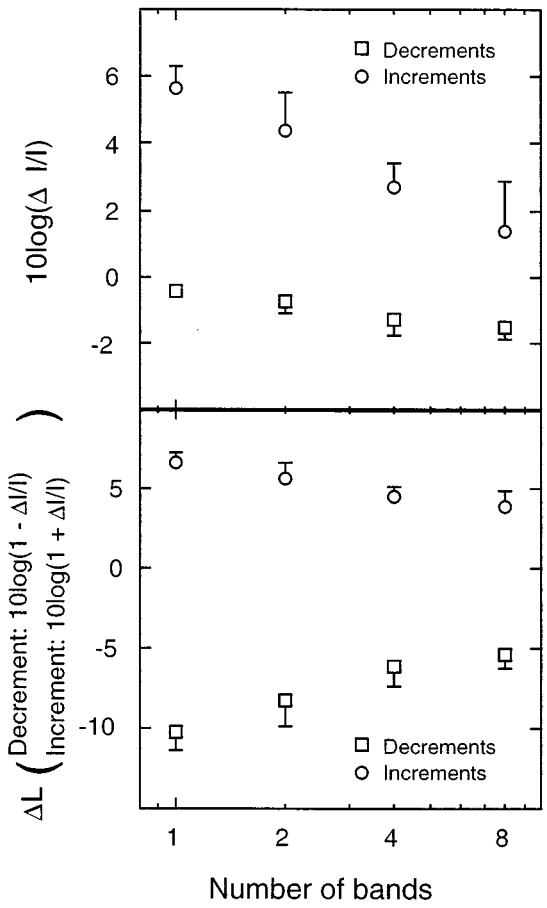


FIG. 4. Group mean results depicting decrement detection (squares) and increment detection (circles). Upper panel shows data expressed in terms of 10 log($\Delta I/I$), with increment data replotted from Fig. 3. Lower panel shows same data expressed in terms of ΔL: for decrements, $\Delta L = 10\log(1-\Delta I/I)$; for increments, $\Delta L = 10\log(1+\Delta I/I)$. Error bars indicate 1 standard deviation.

proportional to ΔL over the entire range of difference limens.[1] In comparisons of increment and decrement detection, there is also a precedent for the predominant use of the ΔL metric (de Boer, 1986; Moore and Peters, 1997). Accordingly, the results of this experiment will be discussed in terms of ΔL. It is apparent from Fig. 4 (lower panel) that, for a single 20-Hz-wide narrow band of noise, a much larger absolute level change is required for decrement detection than for increment detection (10.3 dB versus 6.7 dB). For eight independent narrow bands of noise, the absolute ΔL values are 5.4 and 3.8 dB for the decrements and increments, respectively. The finding that the absolute ΔL for decrement detection is greater than that for increment detection is in line with the data of Moore and Peters (1997) who found that for 200-ms increments and decrements in mid- to high-frequency pure tones, a significantly larger absolute ΔL was required for decrement detection in comparison to increment detection. A similar finding for narrow bands of noise was reported by Hall and Grose (1991).

The results of experiment 2 were interpreted as indicating that intensity discrimination (detection of increments) is a linear function of $\sqrt{N}$. However, a different pattern appears to hold for decrement detection. In order to better understand

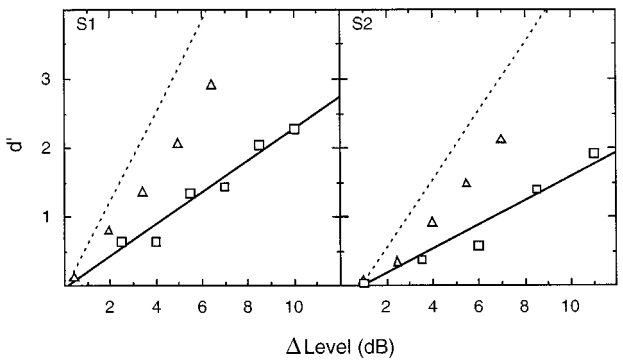


FIG. 5. Sensitivity (d') as a function of ΔL (dB) derived from psychometric functions for decrement detection for two listeners. Squares indicate data for decrements in a single narrow band of noise; triangles indicate data for synchronous decrements in eight narrow bands of noise. See text for explanation of solid and dashed lines.

the behavior underlying decrement detection as a function of the number of stimulus constituents, psychometric functions were generated for two listeners for the single-noise band stimulus and the eight-noise band stimulus. A modified method of constant stimuli was used as implemented by Furukawa and Moore (1996). Blocks comprised of 55 trials were presented which contained five practice trials followed by five sub-blocks of ten constant-decrement trials each. The five sub-blocks spanned a predetermined range of decrement levels, descending from large decrements (high detectability) to small decrements (low detectability). Each listener undertook 20 such blocks per stimulus configuration so that each point on the psychometric function was based on 200 trials. The percent correct scores were converted to d' scores using standard tables (Green and Dai, 1991).

The results for the two listeners are shown in separate panels in Fig. 5. Each panel plots d' against ΔL with number of noise bands (N) as the parameter. Note that d' is plotted on a linear scale while ΔL is plotted on a dB scale. The solid line in each panel is the best-fitting linear regression through the data points for the single-band stimulus (squares) with no constraint for passing through the origin. The dashed line indicates expected performance if d' for eight bands was simply $\sqrt{8}$ better than d' for one band. The results suggest that performance for detecting synchronous intensity decrements in multiple narrow bands improves by less than $\sqrt{N}$ as the number of bands increases. The reason for this effect is not clear. However, it is likely not to be due to differences in sensitivity to decrements as a function of noise band center frequency. Pilot work measuring the detectability of decrements in single 20-Hz-wide bands of noise centered at different points across the frequency range employed in this experiment and presented against the same broadband background noise did not reveal any systematic change in detection threshold with frequency.

In summary, the results of experiment 3 indicate that the detection of decrements in intensity improves by a factor of less than $\sqrt{N}$. This is in contrast to the detection of intensity increments where the $\sqrt{N}$ rule appears to hold.

IV. EXPERIMENT 4. GAP DETECTION AS A FUNCTION OF THE NUMBER OF NOISE BANDS CARRYING THE GAP

The extreme of decrement detection is gap detection. That is, when the level of the antiphasic replica waveform is the same as that of the standard, then perfect cancellation occurs: ΔL is maximal and $10 \log(\Delta I/I) = 0$ dB. Unlike decrement detection, where the duration of the decrement is constant and the procedure titrates along some dimension related to ΔL, in gap detection the level of the decrement is constant and the duration of the decrement is titrated. The issue of interest in experiment 4 is therefore somewhat different from that in experiments 1–3. That is, how does the duration of the just-detectable gap vary as a function of the number of independent noise bands carrying the gap? Several previous studies have addressed aspects of this question (Green and Forrest, 1989; Grose, 1991; Grose and Hall, 1988; Hall *et al.*, 1996). The consensus appears to be that gap threshold is shorter for synchronous gaps in multiple carriers distributed across frequency than for a gap carried by a single frequency-specific carrier. However, most of these studies have not systematically varied the number of carriers and, thus, the function relating gap threshold to carrier number remains unexplored.

A. Method

1. Subjects

The same eight subjects from experiment 1 participated in experiment 4. Each subject received at least 2 h of training on the task, until performance appeared stable; that is, until no systematic improvement in threshold was obvious.

2. Stimuli

The stimuli for experiment 4 were the same eight narrow bands of noise used as maskers in experiment 1 and as standards in experiments 2 and 3. The level of the noise bands remained at 35 dB/Hz and the level of the broadband background noise remained at 15 dB/Hz.

3. Procedure

For each listener, gap threshold was measured for each of the four stimuli. Where stimuli consisted of multiple noise bands, the gap occurred synchronously across all noise bands present. A 3AFC task was again employed where now the target event was the presence of a gap, or interruption, in the stimulus in one of the three observation intervals. The stimulus was gated on for 800 ms in each of the observation intervals. In one of the intervals, chosen at random, the stimulus was gated off and on (Wilsonics BSIT), beginning 400 ms into the interval. The gap was imposed with 20-ms, cosine-squared fall–rise times and the duration of the gap was taken as the time between the initiation of gate closure and the initiation of gate opening. Subjects indicated their responses by means of a microterminal and feedback was provided after each trial. The step size for gap duration was a constant factor of 1.2. A threshold run was terminated after ten reversals in gap duration and the geometric mean of the final six reversal levels was taken as the estimate of threshold

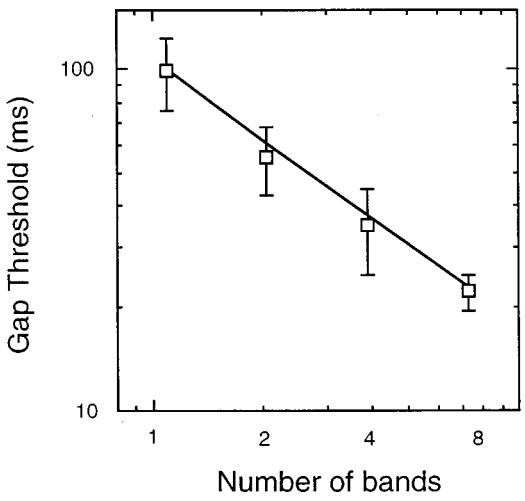


FIG. 6. Group mean results depicting gap threshold as a function of the number of noise bands carrying the gap. Error bars indicate ±1 standard deviation. See text for explanation of solid line.

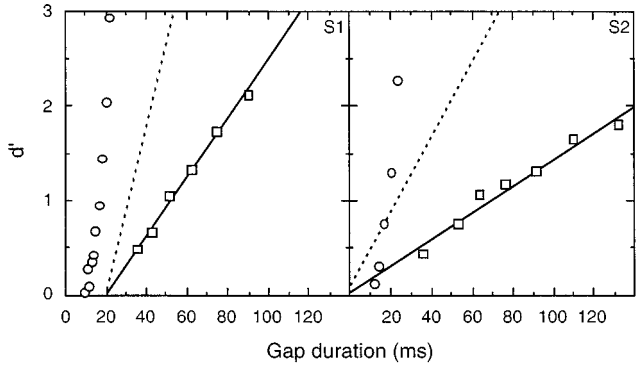


FIG. 7. Sensitivity (d') as a function of gap duration derived from psychometric functions for gap detection for two listeners. Squares indicate data for gaps in a single narrow band of noise; circles indicate data for synchronous gaps in eight narrow bands of noise. See text for explanation of solid and dashed lines.

for that run. For each listener, at least six estimates of threshold were obtained for each condition and the final threshold value was taken as the geometric mean of those six or more threshold estimates.

B. Results and discussion

The group mean results were representative of the eight listeners and are shown in Fig. 6. Gap threshold is plotted against the number of signal components. A repeated measures ANOVA on the logarithmic transform of the data indicated that gap threshold improved significantly as the number of noise bands carrying the gap increased [$F(3,21) = 176.31$; $p < 0.01$]. *Post-hoc* contrasts indicated that gap threshold improved significantly for each doubling of noise band number. The solid line in Fig. 6 is a function whose equation is

$$\log(G_N) = \log(G_1) - \sqrt{2}\log(\sqrt{N}),$$

where G_N is gap threshold for an N-band stimulus and G_1 is gap threshold for a single-band stimulus. The $\sqrt{2}$ term signifies that the improvement in gap threshold with increasing band number is not a linear function of $\sqrt{N}$. Rather, gap threshold improves more steeply than predicted by a simple $\sqrt{N}$ rule. This observation is consistent with the studies of Grose (1991) and Green and Forrest (1989). For example, the Grose (1991) study examined gap detection as a function of the spectral distribution of the noise bands making up the stimulus and found that the improvement in gap threshold when increasing the number of bands from one to five was greater than $\sqrt{5}$. Green and Forrest (1989) also noted that gap threshold appeared to improve with increasing N by a factor greater than would be expected on statistical grounds alone, even when taking into consideration the observation that gap detectability (d') was unlikely to be simply proportional to gap duration.

In order to gain a better understanding of how sensitivity to the presence of a gap depends upon the number of gap carriers, psychometric functions were generated for two listeners in a manner similar to that described for experiment 3. Blocks of 55 trials were presented wherein each block contained five practice trials followed by five sub-blocks of ten constant gap duration trials each. The five sub-blocks spanned a predetermined range of gap durations, descending from long gaps to short gaps. Each point of the psychometric function was based on 200 trials for each listener. Functions were generated for the single-band stimulus and the eight-band stimulus. These data are shown in Fig. 7 where d' is plotted against gap duration for the single-band stimulus (squares) and the eight-band stimulus (circles). Each panel indicates data from a single listener. The solid line in each panel represents the best-fitting linear regression through the data for the single-band stimulus with no constraint for passing through the origin. The dashed line represents expected performance if d' for eight bands was simply $\sqrt{8}$ better than d' for one band. The results suggest that, at least for values of d' not close to zero, performance for detecting synchronous gaps in multiple noise bands improves more steeply with increasing gap duration than would be expected from a simple $\sqrt{N}$ rule. This is in line with the pattern of threshold data seen in Fig. 6.

It is notable that the results of this experiment indicate that gap detection improves more steeply with increasing N than predicted by a $\sqrt{N}$ rule, whereas the results of experiment 3 show that decrement detection improves *less* steeply than $\sqrt{N}$. This contrast is striking because, conceptually, both tasks hinge on the ability to detect a reduction or absence of stimulus energy. A key difference across tasks, however, is that the duration of the change in energy was much longer in the decrement detection task (200 ms) than it was in the gap detection task, suggesting that factors related to temporal integration may have played fundamentally different roles in the two tasks. Further work is required to clarify this issue. For example, a study comparing the detection of shorter-duration decrements and the detection of gaps would be informative. Green and Forrest (1989) have already shown that psychometric functions for partially filled gaps (i.e., constant decrements) are more shallow than psychometric functions for complete gaps. The effect of presentation level is also a matter of consideration. Hall *et al.* (1996), examining the

improvement in gap threshold as the number of bands increased from one to two, found that the degree of improvement appeared to vary as a function of presentation level. At low sensation levels (15 dB SL), gap threshold for two bands was lower than for one band by more than a factor of $\sqrt{2}$. However, at a higher sensation level (30 dB SL), threshold for two bands was roughly a factor of $\sqrt{2}$ lower than for one band. Importantly, for a given presentation level, Hall *et al.* (1996) found that threshold change as a function of the number of carriers was independent of the spectral location of the carriers over a five-octave range.

V. SUMMARY

The purpose of this study was to examine the processing of energy fluctuations which occurred synchronously across relatively independent spectral regions. The results of experiment 1 showed that the detection of multitonal complexes masked by narrow bands of noise was directly related to the square root of the number of signal components $\sqrt{N}$. The function was the same as that observed for broadband masking noise (e.g., Buus *et al.*, 1986; Green, 1958). Thus each signal component appeared to contribute equally and independently to threshold. Experiment 2 extended this to show that, when the signal was a nontonal increment in the level of a noise band carrier, threshold was about 2 dB worse than for a comparable tonal signal. Nevertheless, threshold continued to be a linear function of $\sqrt{N}$, indicating statistical independence of the signal components in the integration process. This was not the case for energy decrements, as shown in experiment 3. The detection of a decrement in energy in one or more noise bands was relatively poorer than was the detection of an increment in energy in the same stimulus. Psychometric functions for decrement detection suggested that performance improved by less than $\sqrt{N}$ with increasing number of stimulus constituents. Experiment 4 examined the detection of energy decrements in the form of imposed silent intervals. As with previous gap detection studies using stimuli presented at low levels (e.g., Green and Forrest, 1989; Grose, 1991; Hall *et al.*, 1996), gap detection was relatively poorer for single narrow bands of noise than for multiple spectrally distributed narrow bands of noise. However, gap detection improved more steeply with increasing N than predicted by a simple $\sqrt{N}$ rule, a finding confirmed with an analysis of psychometric functions for gap detection.

Overall, the results of the first two experiments suggest that, for energy increments, the auditory system integrates information from across the spectrum in a statistically independent manner, at least over the range of frequencies and levels examined here. The detection of relatively long duration energy decrements appears to improve by less than $\sqrt{N}$, whereas the detection of shorter-duration energy gaps appears to improve by more than $\sqrt{N}$. The factors underlying this somewhat paradoxical situation require further investigation.

ACKNOWLEDGMENTS

This work was supported by Grant No. R01 DC01507 from the NIDCD. We thank Dr. Brian Moore, and Dr. Emily Buss, and an anonymous reviewer for their helpful comments on a previous version of this paper. The assistance of Sophie Bonnans, Ann McKinley, and Shelley Warburton in collecting the data is gratefully acknowledged.

[1]Buus and Florentine (1991) expressed ΔL in the pressure domain as $20 \log[(p+\Delta p)/p]$.

Bacon, S. P., and Viemeister, N. F. (**1985**). ''Temporal modulation transfer functions in normal-hearing and hearing-impaired subjects,'' Audiol. **24**, 117–134.

Bos, C. E., and de Boer, E. (**1966**). ''Masking and discrimination,'' J. Acoust. Soc. Am. **39,** 708–715.

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., Schorer, E., Florentine, M., and Zwicker, E. (**1986**). ''Decision rules in detection of simple and complex tones,'' **80,** 1646–1657.

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).

Eddins, D. A., and Green, D. M. (**1995**). ''Temporal integration and temporal resolution,'' in *Hearing*, edited by B. C. J. Moore (Academic, San Diego).

Fitzgibbons, P. J. (**1983**). ''Temporal gap detection in noise as a function of frequency, bandwidth, and level,'' J. Acoust. Soc. Am. **74,** 67–72.

Fletcher, H. (**1940**). ''Auditory Patterns,'' Rev. Mod. Phys. **12,** 47–65.

Florentine, M., and Buus, S. (**1983**). ''Temporal resolution as a function of level and frequency,'' in *Proceedings of the 11th International Congress on Acoustics* **3**, pp. 103–106.

Formby, C., and Muir, K. (**1988**). ''Modulation and gap detection for broadband and filtered noise signals,'' J. Acoust. Soc. Am. **84,** 545–550.

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.

Furukawa, S., and Moore, B. C. J. (**1996**). ''Across-channel processes in frequency modulation detection,'' J. Acoust. Soc. Am. **100,** 2299–2311.

Green, D. M. (**1958**). ''Detection of multiple component signals in noise,'' J. Acoust. Soc. Am. **30,** 904–911.

Green, D. M. (**1988**). *Profile Analysis* (Oxford U. P., Oxford).

Green, D. M., Birdsall, T. G., and Tanner, W. P. (**1957**). ''Signal detection as a function of signal intensity and duration,'' J. Acoust. Soc. Am. **29,** 523–531.

Green, D. M., and Dai, H. (**1991**). ''Probability of being correct with 1 of M orthogonal signals,'' Percept. Psychophys. **49,** 100–101.

Green, D. M., and Forrest, T. G. (**1988**). ''Detection of amplitude modulation and gaps in noise,'' in *Basic Issues in Hearing*, edited by H. Duifhuis, H. P. Wit, and J. P. Horst (Academic, New York).

Green, D. M., and Forrest, T. G. (**1989**). ''Temporal gaps in noise and sinusoids,'' J. Acoust. Soc. Am. **86,** 961–970.

Grose, J. H. (**1991**). ''Gap detection in multiple narrow bands of noise as a function of spectral configuration,'' J. Acoust. Soc. Am. **90,** 3061–3068.

Grose, J. H., Eddins, D. A., and Hall, J. W. (**1989**). ''Gap detection as a function of stimulus bandwidth with fixed high-frequency cutoff in normal-hearing and hearing-impaired listeners,'' J. Acoust. Soc. Am. **86,** 1747–55.

Grose, J. H., and Hall, J. W. (**1988**). ''Gap detection in a narrow-band noise with either a comodulated or noncomodulated flanking band,'' J. Acoust. Soc. Am. Suppl. 1 **83,** S34.

Hall, J. W., and Grose, J. H. (**1991**). ''Relative contributions of envelope maxima and minima to comodulation masking release,'' Quart. J. Exp. Psychol. **43A**, 349–372.

Hall, J. W., Grose, J. H., and Joy, S. (**1996**). ''Gap detection for pairs of noise bands: Effects of stimulus level and frequency separation,'' J. Acoust. Soc. Am. **99,** 1091–1095.

Macmillan, N. (**1971**). ''Detection and recognition of increments and decrements in auditory intensity,'' Percept. Psychophys. **10,** 233–238.

Moore, B. C. J., and Peters, R. (**1997**). ''Detection of increments and decrements in sinusoids as a function of frequency, increment and decrement duration and pedestal duration,'' J. Acoust. Soc. Am. (submitted).

Moore, B. C. J., and Glasberg, B. R. (**1983**), ''Suggested formulae for calculating auditory filter bandwidths and excitation patterns,'' J. Acoust. Soc. Am. **74,** 750–753.
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.
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.
Richards, V. M., and Nekrich, R. D. (**1993**). ''The incorporation of level and level-invariant cues for the detection of a tone added to noise,'' J. Acoust. Soc. Am. **94,** 2560–2574.
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. (**1985**). ''Detection of temporal gaps in bandlimited noise: Effects of variations in bandwidth and signal-to-masker ratio,'' J. Acoust. Soc. Am. **77,** 635–639.
Viemeister, N. F. (**1979**). ''Temporal modulation transfer functions based upon modulation thresholds,'' J. Acoust. Soc. Am. **66,** 1364–1380.
Watson, C. S., and Gengel, R. W. (**1969**). ''Signal duration and frequency in relation to auditory sensitivity,'' J. Acoust. Soc. Am. **46,** 789–797.

The effects of two temporal cues on pitch judgments

Robert P. Carlyon
MRC Applied Psychology Unit, 15 Chaucer Road, Cambridge CB2 2EF, England

(Received 3 October 1996; accepted for publication 3 April 1997)

A series of experiments required listeners to judge which of two sequentially presented pulse trains, bandpass filtered to contain only unresolved harmonics, had the higher pitch. The effects of two temporal cues were explored. The ''common interval'' cue refers to the fact that the interval between any two pulses in each train is an integer multiple of $1/F0$, where $F0$ is the period of the pulse train. The ''mean rate'' cue arises from the fact that the number of pulses in any period of time is proportional to $F0$. It was manipulated, independently of the common interval cue, by removing a proportion of the pulses in each of two trains with different $F0$s. The mean rate cue was removed in one condition of the experiments by deleting more pulses in the stimulus with the higher $F0$; in other conditions, the number of pulses deleted was either independent of $F0$, or was higher for the stimulus with the lower $F0$. The results both of forced-choice and of pitch-matching experiments showed that pitch judgments were affected by the mean rate cue, even though, at least when only a few pulses were deleted, listeners could make consistent judgments using only the common interval cue. Performance in all conditions deteriorated when the number of pulses to be deleted was increased, or when a masking pulse train with a similar spectrum to the targets was mixed with them. Under such circumstances, performance in the absence of the mean rate cue was close to chance. © *1997 Acoustical Society of America.* [S0001-4966(97)04408-1]

PACS numbers: 43.66.Fe, 43.66.Hg [JWH]

INTRODUCTION

To estimate the fundamental frequency ($F0$) of harmonics which are unresolved by the peripheral auditory system, listeners process the temporal features of the outputs of auditory filters that are driven by those harmonics (Hoekstra, 1979; Houtsma and Smurzynski, 1990; Shackleton and Carlyon, 1994). They use this information not only to estimate the pitch of sounds where only the higher (unresolved) harmonics are audible, but also to compare the $F0$s of groups of harmonics, such as the formants of speech, to determine whether or not they arise from the same source (Carlyon *et al.*, 1992; Carlyon, 1994; Carlyon and Shackleton, 1994; Carlyon, 1996). There are a number of models of this form of temporal processing (Moore, 1989; Slaney and Lyon, 1990; Meddis and Hewitt, 1991; Patterson *et al.*, 1995), all of which exploit what will be termed here the ''common interval'' cue. This cue is illustrated schematically in Fig. 1(a) and (b), which shows two pulse trains having different fundamental frequencies ($F0$s). In each case, the time between any two pulses is an integer multiple of a common interval, which is shorter in the sound with the higher $F0$ [Fig. 1(a)] than in the one with the lower $F0$ [Fig. 1(b)]. It is this common interval which is represented along the abscissa of the ''summary autocorrelogram'' (Meddis and Hewitt, 1991) and which is the focus of similar temporal models of pitch (Licklider, 1951; Moore, 1989; Slaney and Lyon, 1990; Patterson *et al.*, 1995).

In 1985, Dobie and Dillier pointed out that there is another difference between the two pulse trains in Fig. 1(a) and (b), which is simply that one contains more pulses than the other. In order to investigate the importance of this ''mean rate'' cue, they deleted some of the pulses in each train, with more pulses being deleted in the stimulus with the higher $F0$, resulting in two pulse trains with identical mean rates [Fig. 1(c) and (d)]. Dobie and Dillier found that listeners still performed above chance on a ''pitch discrimination'' task with these two stimuli, and concluded that listeners could make pitch judgments in the absence of the mean rate cue. However, the pulse trains were not bandpass filtered, and so, as the authors pointed out, the discrimination could have been based on changes in the frequencies of one or more of the lower, resolved, harmonics.[1] The experiments described in the present paper used a similar manipulation to the one described by Dobie and Dillier, with the important difference that the stimuli were bandpass filtered so as to contain only unresolved harmonics. The results show that, although the mean rate cue is not essential for the perception of pitch, it can affect the pitch judgments made by listeners in both forced-choice and pitch-matching experiments. It will be concluded that the mean rate cue can influence the perception of pitch, at least when this is defined loosely (e.g., ''that subjective attribute of sound which admits a rank ordering from low to high''—Ritsma, 1963). The extent to which the data can be interpreted in terms of more specific definitions, such as tone chroma and tone height, will be discussed in Sec. IV B.

In addition to investigating the cues used to judge the pitch of sounds presented in isolation, an important motivation for the present study was to shed light on the potential role of temporal processing in the segregation of concurrent sounds. In a recent study, Carlyon (1996) asked listeners to detect an $F0$ difference (''$\Delta F0$'') imposed on a group of unresolved harmonics which had been mixed with a second, spectrally overlapping, group. (The unresolved harmonics were similar to the bandpass-filtered pulse trains used here, with the minor difference that the harmonics were summed

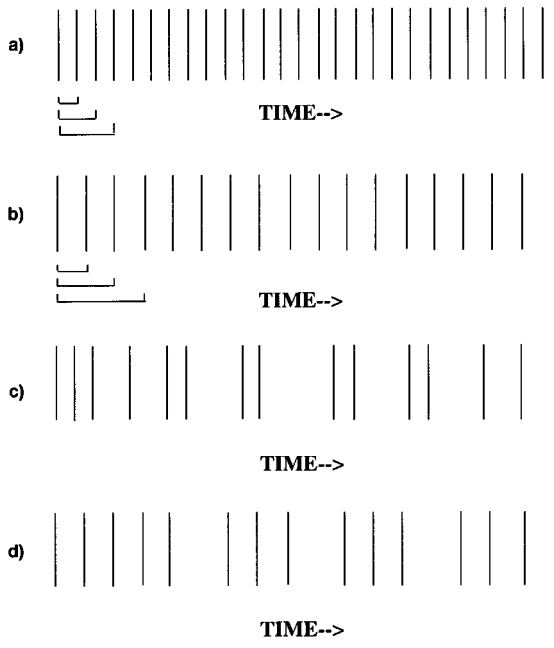


FIG. 1. Parts (a) and (b) show schematic representations of two pulse trains with different $F0$s. The brackets underneath the first few pulses in each train indicate the fact that the time between any two pulses is an integer multiple of a common interval. Parts (c) and (d) show the same pulse trains as in parts (a) and (b), but with some pulses deleted so that the total number of pulses is the same in each stimulus.

TABLE I. Each part of the table shows the $F0$, pulse probability (P), and mean rate for the standard (bold type) and signal (plain type) stimuli for one of the four conditions of experiment 1.

$F0$ (Hz)	P	$R=P\cdot F0$
(a) Baseline		
200.0	**0.707**	**141.4**
208.0	0.707	147.1
216.3	0.707	153.0
234.0	0.707	165.4
273.7	0.707	193.5
(b) Equated		
200.0	**0.707**	**141.4**
208.0	0.680	141.4
216.3	0.654	141.4
234.0	0.604	141.4
273.7	0.517	141.4
(c) Exaggerated		
200.0	**0.517**	**103.4**
208.0	0.537	111.7
216.3	0.559	120.9
234.0	0.604	141.4
273.7	0.707	193.5
(d) Rate		
200.0	**0.517**	**103.4**
200.0	0.559	111.7
200.0	0.604	120.9
200.0	0.707	141.4
200.0	0.968	193.5

in sine rather than in cosine phase.) Although listeners could perform this task, and described the cue they used as a change in pitch, they did not report hearing more than one sound source; rather, they heard the mixture as containing only a single, aperiodic sound. That article concluded that listeners could not use purely temporal cues to segregate two groups of unresolved harmonics, and argued instead that listeners performed the task using a ''composite'' pitch which was conveyed by the average rate at which peaks occurred in the combined waveform. In order to test this conclusion, experiment 3 required listeners to detect an $F0$ difference imposed on a group of unresolved harmonics, masked by another spectrally overlapping group, while controlling for the ''mean rate'' cue.

I. EXPERIMENT 1

A. Method

Experiment 1 consisted of four conditions, termed ''baseline,'' ''equated,'' ''exaggerated,'' and ''rate.'' In each of these, the standard stimulus was a 200-Hz 500-ms pulse train which had been bandpass filtered between 3900 and 5400 Hz. The duration of each pulse was 0.05 ms. In the first three conditions the signal consisted of a pulse train which was passed through the same bandpass filter, but which had an $F0$ of 208, 216, 234, or 274 Hz. Four-point psychometric functions, with 200 trials per point, were measured using the method of constant stimuli as described by Carlyon (1996); each block of trials consisted of ten measures per point, and commenced with six practice trials at the largest $\Delta F0$. On each two-interval forced-choice trial listeners were instructed to identify which stimulus had the higher pitch, and answers were scored as correct when this stimulus corresponded to the one with the higher $F0$. The signals and the method of scoring responses in the fourth (rate) condition will be described later.

For each stimulus presentation, only a given proportion, ''P,'' of the pulses originally present in that pulse train were actually played out, with the remaining pulses being deleted in software. The decision to delete or retain any pulse was taken by generating a random integer between 1 and 10 000, and retaining that pulse if the integer was less than or equal to P multiplied by 10 000. Throughout this article, the term ''$F0$'' is defined as the ''original'' $F0$ of a stimulus even though, strictly speaking, the deletion of pulses rendered the stimuli aperiodic.

In the baseline condition, P was set to 0.707 for the standard and all signals. As shown in Table I(a), this meant that the mean pulse rate ($P\cdot F0$) increased with increasing $F0$; the proportional increase was in fact identical to that occurring in ''conventional'' paradigms, where no pulses are deleted (Hoekstra, 1979; Cullen and Long, 1986). This was not the case in the equated condition, where P equalled 0.707 for the standard but decreased with increasing $F0$ for the signals. As can be seen in the last column of Table I(b), this resulted in the mean pulse rate being the same for the standard and signals. In contrast, the exaggerated condition involved setting $P=0.517$ for the standard, and *increasing* P for signals with increasing $F0$s [Table I(c)]. In this condition, then, the increase in mean pulse rate with increasing $F0$ was greater than in a conventional paradigm. Finally, in the rate condition, the $F0$ of the standard *and the signals* was 200 Hz, with the signals differing from the standard

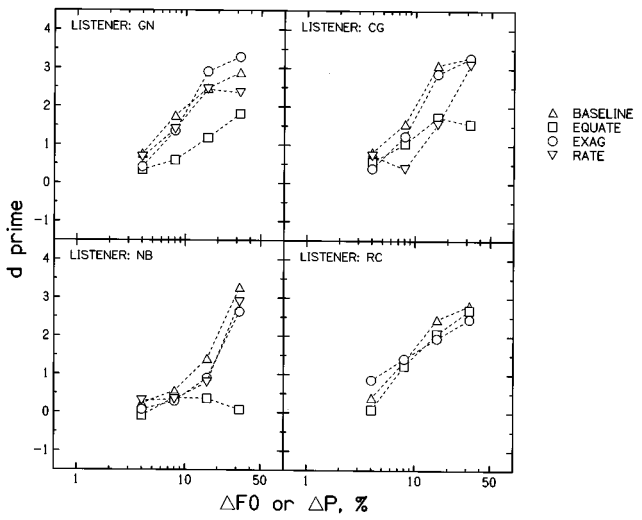


FIG. 2. Sensitivity (d') measured in the four conditions of experiment 1. For the baseline (triangles), equated (squares), and exaggerated (circles) conditions, the abscissa shows the percentage difference in $F0$ between the standard and each signal. For the rate condition, the abscissa shows the percentage difference in pulse probability between the standard and the signals.

only by an increase in the value of P [Table I(d)]. In this condition, responses were scored as correct whenever the listener identified the stimulus with the higher mean rate as having the higher pitch.

An important feature of the experimental design was that the conditions were interleaved, so that, on a given block of trials, listeners did not know which condition they were being tested on. This was done in order to encourage them to identify the signal using the same perceptual attribute in all conditions. Because the decision to include the rate condition was taken after listener RC (the author) had been tested on the other three conditions, the policy of not allowing listeners to know which condition they were being tested on required the rate condition to be omitted for this listener.

The pulse trains were played out of a 12-bit DAC (CED 1401) at a sampling rate of 20 000 Hz, and were passed through an antialiasing filter (Kemo VBF 25.01, cutoff=8600 Hz, attenuation rate 135 dB/oct) before bandpass filtering (one high-pass and one low-pass Kemo VBF 25.03 filters in series, attenuation rate 48 dB/oct). The bandpass filter had 3-dB-down points of 3900 and 5300 Hz. The pulse trains were then attenuated (TDT PA4) and fed to one input of a headphone amplifier. The attenuation was chosen so that an ''intact'' pulse train ($P=1.0$), with an $F0$ of 200 Hz, would produce harmonics in the passband of the filter with levels of 45 dB SPL each. The amplitude of the pulses in the stimuli actually used in the experiments were equal to that of this calibrating stimulus. A continuous pink noise, with a spectrum level of 9 dB SPL at 4000 Hz, was connected to another input of the headphone amplifier. Both the pulse trains and the noise were presented to one earpiece of a Sennheiser HD414 headset. Four listeners, each with absolute thresholds within 15 dB of the 1969 ANSI standard at all audiometric frequencies, listened in a double-walled sound-insulating booth.

B. Predictions

The relative levels of performance in the four conditions of experiment 1 should depend on the effect of mean rate cues on listeners' judgments of pitch. A strong prediction is that, if listeners use only the mean rate cue, then performance in the equated condition should be at chance, but that it should be above chance in the other three conditions. A different prediction arises from the scenario where listeners can make reliable pitch judgments based on the common interval cue alone, but where these judgments are partly influenced by the mean rate cue. In this case, performance in the equated condition should be above chance, but worse than in the baseline or exaggerated conditions. In addition, performance on the rate condition should also be above chance. Finally, if the mean rate cue has no effect, performance should be roughly similar in the equated, baseline, and exaggerated conditions, but that in the rate condition should be at chance.

C. Results

The results of experiment 1 for the four listeners are shown in the four panels of Fig. 2. The abscissa shows the percentage $\Delta F0$ between the standard and the various signals for the baseline, equated, and exaggerated conditions; for the rate condition, it shows the difference in rate between the standard and signals. In each case, the ordinate shows sensitivity measured as d'.

For listener GN (top left), performance in the equated condition (squares) is above chance, but lower than in the baseline (triangles) or exaggerated (circles) conditions. This suggests that his pitch judgments are partly, but not completely, influenced by the mean rate cue; further evidence for a role for the mean rate cue comes from his above-chance performance in the rate condition (inverted triangles). Listener CG shows a similar pattern of results. The results are also similar for listener NB, with the exception that he performs at chance in the equated condition, suggesting that his judgments are entirely dominated by the mean rate cue. The major difference comes from the data of listener RC, whose performance in the equated condition is similar to that in the baseline and exaggerated conditions. This listener differed from the other three in that there was no evidence for an effect of the mean rate cue.

II. EXPERIMENTS 2 AND 3: EFFECT OF A MASKING PULSE TRAIN AND OF REDUCED PULSE PROBABILITY

A. Rationale

In one of the experiments described by Carlyon (1996), listeners were required to detect a $\Delta F0$ imposed on one group of unresolved harmonics that had been mixed with a second, masking group of harmonics in the same frequency region. It was concluded that, in order to perform this task, listeners relied entirely on changes in the average rate at which peaks occurred in the composite waveform of the masker and target complexes. Experiment 2 therefore tested the hypothesis that the mean rate cue would have a more

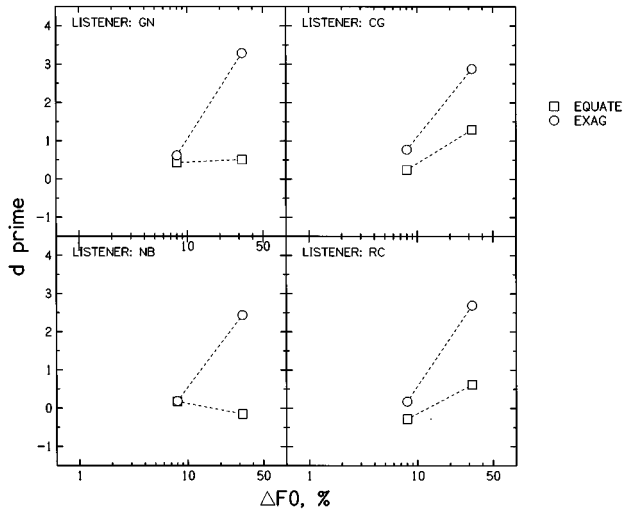


FIG. 3. Results of the equated (squares) and exaggerated (circles) conditions of experiment 2. The abscissa shows the percentage difference in $F0$ between the standard and each signal.

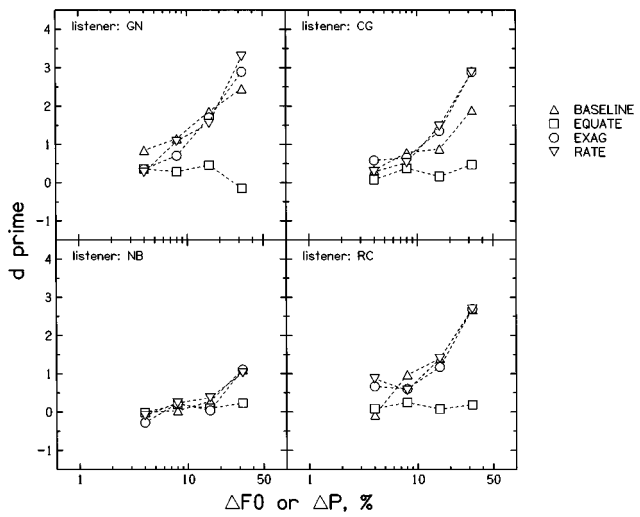


FIG. 4. Sensitivity (d') measured in the four conditions of experiment 3. For the baseline (triangles), equated (squares), and exaggerated (circles) conditions, the abscissa shows the percentage difference in $F0$ between the standard and each signal. For the rate condition, the abscissa shows the percentage difference in pulse probability between the standard and the signals.

consistent effect on performance when the standard and signal pulse trains were mixed with a masking pulse train, compared to the rather variable effects seen in experiment 1. Experiment 3 measured the effect of another manipulation designed to degrade the pitch strength of the stimuli, which was simply to delete a higher proportion of pulses in all conditions.

B. Method

The method used for experiment 2 was based on that for experiment 1. Only the equated and exaggerated conditions were run, and only two signal $F0$s, 217 and 274 Hz, were tested. The main difference was that all stimuli were mixed with a 234-Hz pulse train, having a pulse probability (P) of 0.604, which was gated on and off with the standard and signal pulse trains. It was played out of a second DAC, bandpass filtered and attenuated in a manner identical to the standard and signal trains, and fed into one input of the headphone amplifier. The level of each pulse in this masking stimulus was the same as that in the standard and signals.

Experiment 3 included all conditions of experiment 1, the only difference being that all pulse probabilities (both in the standard and signal intervals) were divided by 1.414. Unlike in experiment 1, listener RC was tested on all four conditions.

C. Results

The results of experiment 2 are shown in Fig. 3. It can be seen that performance was worse in the equated than in the exaggerated condition for all four listeners, including RC, who had shown no such difference in experiment 1. Indeed, performance in the equated condition was close to chance, with the possible exception of listener CG at a $\Delta F0$ of 37%. This suggests that, in the presence of a spectrally overlapping masker, all listeners used the mean rate cue to make pitch judgments (cf. Carlyon, 1996). In order to determine whether the effect of the mean rate cue was greater in experiment 2 than in experiment 1, the data points which were measured in the two experiments were submitted to a three-way (masker presence/absence$\times$condition $\times\Delta F0$) ANOVA. Although the ANOVA revealed significant effects of condition [$F(1,3)=44.1$; $p<0.01$] and of $\Delta F0$ [$F(1,3)=130.0$; $p<0.01$], the effect of masker just failed to reach significance [$F(1,3)=8.6$; $p=0.06$], and, more importantly, did not interact with condition [$F(1,3)=2.5$; $p=0.21$]. This means that, although listener RC only showed a difference between the two conditions when a masking train was present (experiment 2), this interaction was not observed when the data from all four listeners were taken into account.

Figure 4 shows that the effect of reducing pulse probability was, roughly speaking, similar to that of adding a masker. All four listeners showed very poor performance in the equated condition, and rather better performance in the baseline and exaggerated conditions. To compare the effect of the mean rate cue to that in experiment 1, the data from the baseline, equated, and exaggerated conditions from the two experiments were submitted to a three-way (pulse probability$\times$condition$\times\Delta F0$) ANOVA.[2] All main effects were highly significant [pulse probability: $F(1,3)=77.1$; $p<0.01$; condition: $F(2,6)=36.5$, $p<0.01$; $\Delta F0$: $F(3,9)=87.7$, $p<0.01$]. In addition, $\Delta F0$ interacted significantly both with pulse probability [$F(3,9)=12.1$, $p<0.01$] and with condition [$F(6,18)=11.2$, $p<0.01$]. Once more, however, there was no statistical evidence that the effect of mean rate differed from that in experiment 1, as demonstrated by a lack of significant interaction between pulse probability and condition [$F(2,6)=0.3$, $p=0.73$]. This may seem surprising, given the fact that no listener performed above chance in the equated condition of experiment 3. However, this must be attributed to the marked reduction in sensitivity in all conditions of experiment 3, compared to the corresponding conditions of experiment 1; because performance was initially high in the baseline and exaggerated

conditions of experiment 1, reducing pulse probability did not bring performance down to the chance levels observed for the equated condition. The exception to this overall conclusion is again listener RC, whose performance in the equated condition was worse than in the baseline and exaggerated conditions only in experiment 3. Again, we must conclude that the interaction observed for RC did not generalize to the other three listeners.

The overall reduction in performance as pulse probability was decreased from experiment 1 to experiment 3 may explain one further aspect of the data. Although both experiments showed that, relative to the baseline condition, performance was reduced when the mean rate cue was removed (equated condition), performance did not improve when the cue was enhanced (exaggerated condition). A comparison of parts (a) and (c) of Table I reveals that the pulse probabilities were lower, overall, in the exaggerated than in the baseline condition, and this may have counteracted any advantage to be gained from exaggerating the difference in mean rate between the standard and signal. The confounding influence of overall pulse probability can be avoided when comparing performance in the equated and exaggerated conditions. At a signal $F0$ of 274 Hz, the pulse probabilities in the equated condition for the standard and signal were 0.707 and 0.517, respectively, with these values being simply swapped for the exaggerated condition [compare Table I(b) and (c)]. Inspection of the data in Figs. 2–4 shows that at this signal $F0$, which corresponds to a $\Delta F0$ of 37%, performance is substantially better in the exaggerated condition (circles) than in the equated condition (squares). This comparison, then, shows that there is an effect of the mean rate cue even when overall pulse probability is controlled for.

III. EXPERIMENT 4: PITCH MATCHES

A. Rationale and method

Experiments 1–3 provided strong evidence that pitch judgments can be influenced by the mean rate cue. However, those experiments used pulse trains with fairly low pulse probabilities: even in experiment 1, the highest value of P on any trial in the baseline, equated, or exaggerated conditions was 0.707, and all except the baseline condition included trials where either the standard or signal had fewer than 52% of the original pulses remaining (Table I). This raises the possibility that the usefulness of the ''mean rate'' cue is restricted to the case where a fairly large proportion of pulses has been deleted, rather than being generalizable to ''intact'' stimuli. Two aims of experiment 4 were to determine whether the mean rate cue affects the pitch of pulse trains when a high proportion of the original pulses has been retained, and whether the size of any such effect varies with the value of this proportion. A third aim was to study the effect of the mean rate cue using a different technique, pitch matching, which is commonly used in the study of pitch perception. Finally, the experiment allowed us to quantify the effect of a given change in pulse probability on the perceived $F0$ of a sound.

There were three conditions in the experiment, differing only in the lowest pulse probability used for any match in that condition. This value, termed $P_{\min}$, was either 0.55, 0.7, or 0.85. Half of the pitch matches in each condition started with the listener being presented with a 500-ms 200-Hz bandpass-filtered pulse train, generated in the same way as in experiment 1, and with the pulse probability for this ''fixed'' stimulus, P_{fix}, equal to $P_{\min}$. This was followed 500 ms later by a second ''variable'' pulse train, with an $F0$ drawn at random from a distribution which was rectangular on a log axis and which ranged from 154–260 Hz. Its pulse probability, P_{var}, was 10% higher than $P_{\min}$. The listener could then adjust the $F0$ of the second stimulus to be presented on the next trial, by pressing one of four buttons on the response box: the outer two buttons allowed the $F0$ to be multiplied or divided by a factor of 1.1, with a smaller factor of 1.03 for the inner two buttons. Whichever button was pressed, the stimulus parameters for the first pulse train in the next trial remained fixed ($F0=200$ Hz, $P_{\text{fix}}=P_{\min}$), as did the pulse probability for the second, variable, stimulus. A different random set of pulses to be retained was calculated afresh for each pulse train. Listeners continued to adjust the $F0$ of the variable stimulus until they were satisfied that the two stimuli had the same pitch, and then terminated the match by pressing a fifth button; they were encouraged to ''bracket'' the match before doing so.

The other half of the pitch matches for each condition were similar, except that the pulse probability for the fixed stimulus was 10% higher than that for the variable stimulus, instead of *vice versa*. This was done in order to separate the effects of pulse probability from any bias listeners may have had towards adjusting the second stimulus of any pair to either a lower or higher $F0$ than the first.

In all other respects the method of stimulus generation, including the continuous pink noise, was the same as in experiment 1. Three listeners took part, and made nine matches in each half of the three conditions. One listener, CL, had not taken part in experiments 1–3.

B. Results

The $F0$ to which the second, variable stimulus was adjusted is shown in Fig. 5, each panel of which shows the data for three listeners at a single value of $P_{\min}$. The adjusted $F0$ was greater when the variable stimulus had a lower pulse probability than the fixed stimulus (solid bars) than vice versa (hashed bars), consistent with a reduction in P lowering perceived pitch, and with listeners compensating for this by increasing the $F0$ of the variable stimulus. A two-way ANOVA (overall probability×stimulus with higher or lower P varying) confirmed the statistical significance of this finding [$F(1,2)=29.8$; $p<0.05$]. The figure also shows that on several occasions listeners displayed a bias towards adjusting the $F0$ of the variable stimulus to be either lower (e.g., listener CL, $p=0.55$) or higher (e.g., GN, $p=0.55$) than the 200-Hz fixed sound, regardless of which of the fixed or variable sounds had the higher pulse probability.

Table II shows, for each listener and condition, the square root of the ratio between the pitch matches for $P_{\text{fix}}=1.1\cdot P_{\text{var}}$, and that when $P_{\text{var}}=1.1\cdot P_{\text{fix}}$. This value, which represents the change in perceived $F0$ produced by a 10%

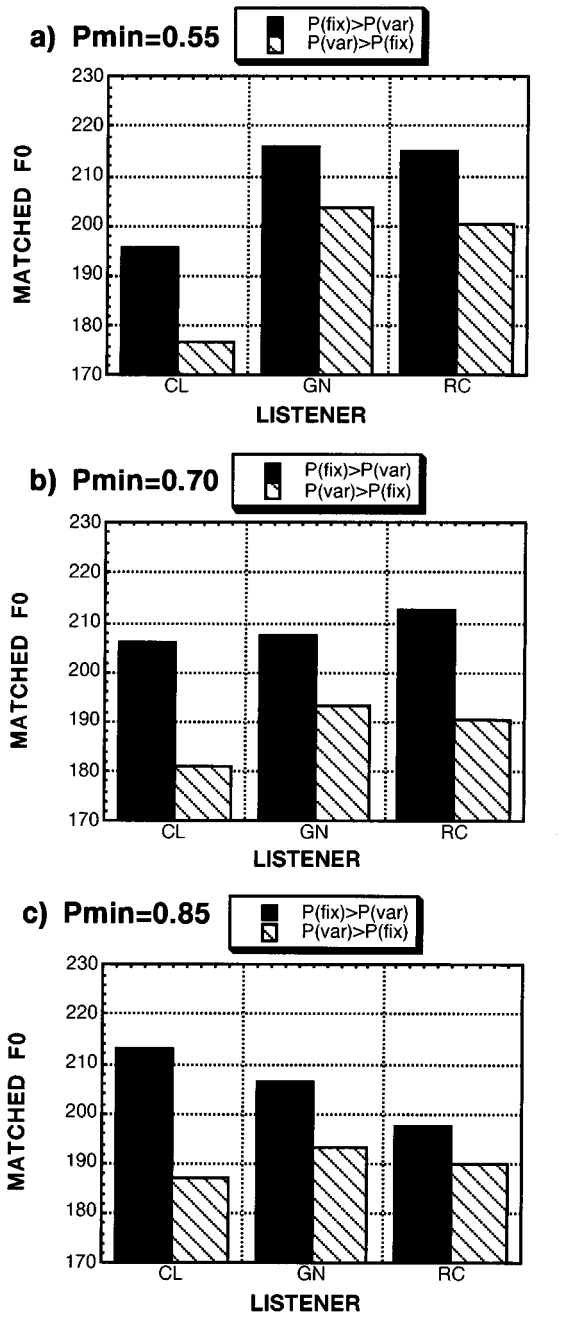


FIG. 5. Pitch matches measured in experiment 4 for $P_{\min}=(a)\,0.55$, (b) 0.70, and (c) 0.85. Solid bars show the $F0$ to which the variable stimulus was adjusted for $P_{\text{fix}}=1.1\cdot P_{\text{var}}$, whereas shaded bars are for $P_{\text{var}}=1.1\cdot P_{\text{fix}}$.

change in P, is on average 1.04, 1.05, and 1.04 for $P_{\min}$ equal to 0.55, 0.70, and 0.85, respectively. Thus, a 10% change in pulse probability produces an approximately 4%–5% change in perceived $F0$, a relationship which does not vary markedly with $P_{\min}$.[3] Table II also shows that the standard deviation of the pitch matches increased with decreasing values of $P_{\min}$. This is almost certainly due to a decrease in pitch strength as more pulses are deleted, and is consistent with the overall drop in performance as pulse probability was lowered from experiment 1 to experiment 3. The mean standard deviation of 4% for $P_{\min}=0.85$ can be compared to that observed for ''intact'' groups of unresolved harmonics by Shackleton and Carlyon (1994).[4] They reported that, for a group of harmonics with an $F0$ of 250 Hz and filtered in the same way as the present stimuli, the standard deviation was approximately 6%, slightly larger than that observed here. This suggests that the pitch percept obtained with $P_{\min}=0.85$ is about as accurate as that obtained for unresolved harmonics with no pulses missing.

TABLE II. Summary of the results from experiment 4. For each listener, the numbers in bold type show the square root of the ratio between the adjusted $F0$ when $P_{\text{fix}}=1.1\cdot P_{\text{var}}$, and that when $P_{\text{var}}=1.1\cdot P_{\text{fix}}$. This value reflects the change in perceived $F0$ produced by a 10% change in P. The numbers in plain type show the standard deviation of the matches in each condition (the square root of the mean of the variances in the $P_{\text{fix}}=1.1\cdot P_{\text{var}}$ and $P_{\text{var}}=1.1\cdot P_{\text{fix}}$ subconditions), converted to a percentage of 200 Hz. Mean data are shown in the last two rows.

		$P_{\min}=0.55$	$P_{\min}=0.70$	$P_{\min}=0.85$
CL	$\sqrt{\text{Ratio}}$	**1.051**	**1.066**	**1.068**
	s.d. (%)	8.84	6.08	3.13
GN	$\sqrt{\text{Ratio}}$	**1.030**	**1.035**	**1.034**
	s.d. (%)	9.70	6.95	6.52
RC	$\sqrt{\text{Ratio}}$	**1.036**	**1.056**	**1.019**
	s.d. (%)	6.97	4.45	2.56
Avg.	$\sqrt{\text{Ratio}}$	**1.039**	**1.052**	**1.040**
	s.d. (%)	8.50	5.83	4.07

IV. DISCUSSION

A. Perceptual correlates of pulse probability

The experiments described here have shown that pulse probability can affect the results of both forced-choice and pitch-matching experiments, and that listeners consistently judge stimuli with lower probabilities to have lower pitches. Before concluding that the results demonstrate a genuine effect of the mean rate cue on pitch, it is worth considering alternative explanations for the results. One possibility is that, because the level of each pulse was constant throughout the experiments, the apparent increase in pitch with increasing mean rate may have been mediated by changes in loudness. However, although it is well known that pitch can vary with loudness, such effects are usually very small, and even a 60-dB increase in the level of a 4000-Hz tone fails to produce variations in pitch as large as the 4% observed in experiment 4 (Burns, 1982). In addition, and in contrast to the consistent direction of the effect observed here, the effects of level on pitch vary in direction across listeners (Burns, 1982).

A variant of the above explanation is that, although loudness differences between our stimuli did not substantially affect pitch, listeners learned to associate these differences with variations in a ''genuine'' (i.e., common rate) pitch cue. This may have occurred because mean rate was correlated with $F0$ in the baseline and exaggerated conditions of experiment 1, and listeners may have generalized this association to other conditions and experiments. This in turn might explain why they could perform above chance in the rate condition, in which the common interval cue was

absent, but in which loudness increased with increasing pulse probability. However, it seems unlikely that it could account for the fact, apparent in Figs. 2 and 4, that performance was generally better in the rate condition than in the equated condition. This would require listeners to have placed a greater weight on the ''spurious'' loudness cue than on the ''genuine'' common interval cue, despite having only adopted the former cue by its association with the latter, and despite instructions to respond to pitch rather than to loudness. The explanation also fails to account for the effect of mean rate on the pitch matches made in experiment 4 by listener CL, as he had not taken part in experiments 1–3, and could therefore not have learned to associate pitch with loudness in the ''baseline'' and ''exaggerated'' conditions.

Another possibility is that listeners' responses were affected by the sensation of roughness which occurs when some of the pulses are removed from a periodic pulse train, and that they learned to associate an increase in roughness with a decrease in pitch. This too seems unlikely, given that roughness (defined here as being inversely proportional to pulse probability) was correlated with the common interval cue only in the exaggerated condition of experiment 1. Rather, although we cannot rule out the possibility that some form of bias had a small effect on our results, it seems reasonable to conclude that the substantial and consistent effects observed here reflect an influence of the mean rate cue on at least loose definitions of pitch (e.g., ''that subjective attribute of sound which admits a rank ordering from low to high''—Ritsma, 1963).

B. Nature of pitch changes

1. Tone height and chroma

Traditional musical representations of pitch are two-dimensional, and often take the form of a helix, with tone chroma varying along the circular dimension, and tone height varying along the length of the helix. However, neither the forced-choice nor pitch-matching techniques used here and elsewhere distinguish between these two dimensions, and so it is worth considering the extent to which the present results can be interpreted as an effect of mean rate on chroma, on tone height, or on both.

The strongest evidence that temporal processing can mediate the perception of chroma comes from the finding that listeners can reproduce musical intervals conveyed by the repetition rate of amplitude modulated noise (Burns and Viemeister, 1981), by the $F0$ of unresolved harmonics (Houtsma and Smurzynski, 1990), and by the rate of impulses applied to single channels of cochlear implants (Pijl and Schwarz, 1995). In the context of the present experiments, this raises the obvious question of whether the mean rate cue can by itself support musical interval recognition and/or production. Unfortunately, we do not have this information, and preliminary considerations suggest that it may be hard to obtain. The strength of the pitch percept decreases markedly as pulse probability is reduced, and, given that a 10% decrease in pulse probability changes the perceived $F0$ by only 4%–5%, any interval recognition task is likely to involve at least one sound with a low pulse probability and therefore a weak pitch.[5] In this regard it is worth noting that the same obstacle confronts attempts to isolate the ''common interval'' cue, as such attempts would also require the deletion of some pulses.

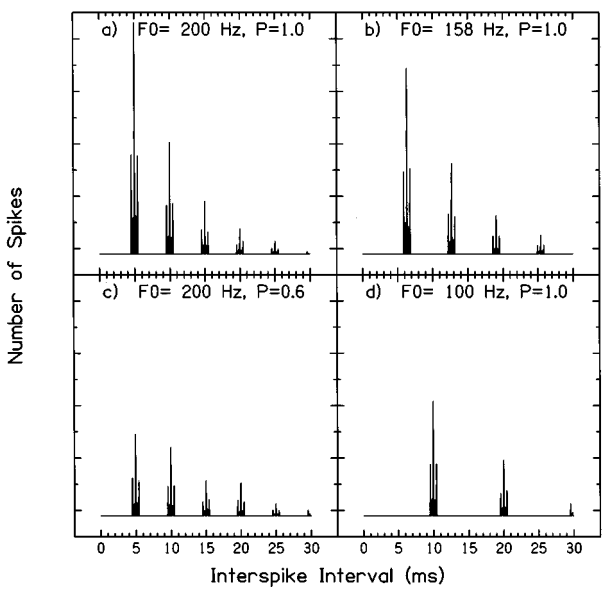


FIG. 6. Interspike interval histograms (ISIH) simulated using the response of an auditory filter centered on 4500 Hz to pulse trains filtered in the same way as in the experiments. Each part shows the result of applying the random algorithm described in the text to a single stimulus, 500 times. Part (a) $F0=200$ Hz, $P=1.0$; (b) $F0=158$ Hz, $P=1.0$; (c), $F0=200$ Hz, $P=0.6$; (d) $F0=100$ Hz, $P=1.0$.

There is also evidence that tone height can be affected by both spectral and temporal cues, and that it can be varied in a continuous manner, independently of chroma. Patterson *et al.* (1993) varied the attenuation applied to the odd harmonics of a complex tone, and observed a gradual shift in the perceived tone height with increasing attenuation. This is not surprising because, as they observed, an infinite attenuation simply doubles the $F0$ of the complex, resulting in an increase in pitch of one octave. A similar increase can be produced in the temporal domain, by shifting the phase of alternate components of a sine- or cosine-phase complex by 90°, provided that the components are unresolved by the peripheral auditory system (Patterson *et al.*, 1993; Shackleton and Carlyon, 1994). The important point about both these manipulations is that one would not expect listeners to make musical interval judgments to anything other than the octave or fundamental, even for intermediate attenuations or phase shifts.

2. Interspike interval histograms

As mentioned above, the procedures used in the present study do not distinguish between the two musical dimensions of pitch, and so we have no firm evidence that one or other dimension was primarily affected by pulse probability. However, some insight may be gained by considering the ways in which the manipulations used here might affect neural representations of temporal information. Figure 6(a) shows a simple simulation of an auditory-nerve interspike interval histogram (ISIH), derived from the rectified output of a ''gammatone'' auditory filter centered on 4500 Hz (Patterson

et al., 1988) in response to a 400-ms pulse train with $F0 = 200$ Hz and $P = 1.0$. The simulation was implemented by assuming that a spike was generated every time the output exceeded a criterion amount, which varied randomly by $\pm 25\%$ from sample to sample of the digitized (20 000 Hz) filter output, subject to an absolute ''refractory time'' of 2 ms. The resulting ISIH shows a large peak at the period of 5 ms, and several small peaks at integer multiples of that period. When the $F0$ is reduced to 158 Hz [Fig. 6(b)], the locations of these peaks shift to multiples of the new period. In contrast, when the pulse probability is reduced to 0.6 [Fig. 6(c)] while keeping $F0$ equal to 200 Hz, the peaks do not shift, but, in addition to an overall decrease in the number of spikes, change height relative to each other. This is in some ways similar to the effect of reducing the $F0$ by an octave [Fig. 6(d)]: in each case, none of the peaks change position, but the one at 10 ms increases in height relative to that at 5 ms. The main difference between the two manipulations is that reducing pulse probability [Fig. 6(c)] leaves peaks present at odd multiples of the original period, whereas the octave shift [Fig. 6(d)] does not. Nevertheless, the similarity between parts (c) and (d) of Fig. 6 tentatively suggests that changing P may cause listeners to make a pitch judgment based on a dimension akin to the octave (i.e., tone height), rather than to ''chroma.'' Thus, in terms of ISIHs, the signals in the equated conditions of experiments 1 and 3 would have differed from the standard in two ways: the individual peaks in the ISIH would have occurred at shorter delays than those in the standard, but relative heights of the peaks would have changed in such a way as to favor longer delays. If listeners were using some combination of these changes when judging pitch, then they would have less reliably identified the signal as having a higher pitch than the standard in the equated condition than they would have in the baseline and exaggerated conditions. Such a pattern of results can be observed in Figs. 2 and 4.

C. Pulse probability and sound pressure level

Further understanding of the way the auditory system processes temporal information can be obtained by comparing the effects on the firing patterns of auditory nerve fibers of reducing P, with those of reducing sound pressure level. Provided that the interpulse interval $(1/F0)$ is long enough to obviate the effects of refractoriness and of adaptation, the firing pattern of an individual auditory nerve fiber can be modeled as a Poisson process (Siebert, 1968). A feature of this process is that the probability of any one spike occurring is independent of the occurrence of earlier spikes, a characteristic shared by our generation of random pulses in the stimulus. Thus, for a single neuron, the effect of reducing level should be similar to that of reducing P: fewer spikes will occur, with the ''missing'' spikes distributed at random throughout the spike train. Informal observations reveal that the two manipulations produce very different percepts, even when $1/F0$ is long enough (e.g., 10 ms) to rule out refractory effects. Therefore, the firing patterns of several fibers must be combined in some way. Furthermore, this combination must take place *before* the initial transformation (e.g., ISIH, autocorrelogram) is made, rather than individual transformations being applied to single auditory nerve fibers. In the latter case, the auditory system would sum a set of representations each of which did not distinguish between a change in level and a change in P, and so the total ''summary representation'' would not be able to make this distinction.

There are two ways in which information from different auditory neurons can be combined in order to derive a peripheral representation of pitch which can discriminate between level and pulse probability. The firing patterns of neurons with similar characteristic frequencies (CFs), but which innervate different inner hair cells, are independent of each other, and the composite pattern from a number of such fibers will differ between the two manipulations. A reduction in level will produce some ''missing spikes'' in individual neurons, but the absence of these spikes will occur at different times for different neurons, and so every pulse will produce some activity in the composite pattern. In contrast, a reduction in P will result in no fibers firing at those times when a pulse has been deleted from the train, and so the composite pattern will contain ''silent'' periods similar to those in the stimulus.

In principle, a similar manipulation could be performed for fibers with widely different CFs. Evidence against this idea comes from experiments which show that temporal information from different channels cannot be combined into a single percept: for example, listeners cannot discriminate between AM and FM tones when the components present in each tone are separated by more than a critical band (Zwicker, 1952). Nevertheless, listeners may be able to use across-channel cues by correlating the outputs of different channels (Hall *et al.*, 1984; McFadden, 1987; Richards, 1987; Carlyon, 1994), as this correlation will be greater for a train with a low pulse probability than for one with a low level.

D. Temporal cues for concurrent sound segregation

As discussed in the Introduction, it has recently been argued (Carlyon, 1996) that listeners cannot exploit $F0$ differences in order to segregate two spectrally overlapping groups of unresolved harmonics. Those articles concluded that, although listeners can detect an increase in the $F0$ of one of the groups, they do so by exploiting an increase in the ''mean rate'' of the composite stimulus. This conclusion is supported by the results of experiment 2, which showed that, when the mean rate cue was removed, performance dropped to chance. Thus, at least for the monotically presented stimuli used in this laboratory, we can find no evidence that $F0$-based segregation is aided by a within-channel mechanism such as de Cheveigné's (1993) ''neural cancellation filter.'' This is not to suggest that temporal mechanisms do not play a role in concurrent sound segregation: indeed, quite the opposite is suggested by Hartmann's (1988) finding that sensitivity to the mistuning of a resolved harmonic depends on its phase. However, the data presented by Carlyon (1996) and extended here do indicate that, although temporal processes may guide the selection of channels associated with a particular source (cf. Assmann and Summerfield, 1990), they are not capable of extracting multiple $F0$s from within a single channel.

ACKNOWLEDGMENTS

I thank Hedwig Gockel, Alain de Cheveigné, Roy Patterson, Ed Burns, and an anonymous reviewer for constructive comments on earlier versions of this paper. Lutz Wiegrebe wrote the computer program which transformed the gammatone filter outputs to the ISIHs shown in Fig. 6, and Jay Datta helped with data collection.

[1]A similar finding was obtained when the pulses were presented to cochlear implant listeners by stimulating a single electrode, thereby eliminating spectral cues. This suggests that implantees can perceive pitch in the absence of mean rate cues, but does not show that such cues have no effect at all on temporally based pitch perception. An additional reason why Dobie and Dillier may not have ruled out the use of mean rate cues to pitch is that they did not state whether or not listeners were provided with correct-answer feedback. In order to equate the mean rates of two pulse trains having different $F0$s, they had to delete more pulses from one of the trains, which would have resulted in it having a ''rougher'' or more ''crackle-like'' quality. If listeners were provided with feedback, then they might have exploited this cue, thereby achieving above-chance performance without necessarily perceiving the sound with the higher $F0$ as having the higher pitch.

[2]The rate condition was omitted from the analysis because one listener did not take part in that condition in experiment 1.

[3]Pilot data from a fourth listener, who was a professional psychoacoustician and who listened diotically, showed an effect of mean rate only at low pulse probabilities.

[4]Shackleton and Carlyon's stimuli were passed through the same filters as used here, and had a level 5 dB higher than that of the present stimuli (when $P=1.0$). This level difference was counteracted by the pink noise they used having a level 5 dB higher than that presented here. The only major difference was that their stimuli were frequency modulated by $\pm 5\%$ at a rate of 2.5 Hz. However, data from a different experiment in that study (their Fig. 4) show that this modulation was not detectable, and so is unlikely to have influenced the variability of the measured pitch matches. A further difference is that their harmonics were summed in sine phase, rather than the cosine relationship that exists between the harmonics of a pulse train. Both phase relationships result in highly peaked waveforms, which are unlikely to differ markedly in pitch strength.

[5]This was confirmed by informal attempts to have the author (who is not musically trained) and a musically trained listener identify simple melodies from a small closed set, using stimuli similar to those in the present study. Although the task was easy when no pulses were deleted, reducing the pulse probability to 0.7 (for all notes, as in our ''baseline'' condition) made the task much harder, and we were unable to observe performance that was consistently above chance. This was also true for melodies in which intervals were conveyed by only the ''common interval'' cue (cf. our equated condition), or by only the mean rate cue (cf. the rate only condition).

ANSI. (**1969**). ANSI S3.6-1969, ''Specifications for audiometers'' (American National Standards Institute, New York).

Assmann, P., and Summerfield, Q. (**1990**). ''Modeling the perception of concurrent vowels: Vowels with different fundamental frequencies,'' J. Acoust. Soc. Am. **88,** 680–697.

Burns, E. M. (**1982**). ''Pure tone pitch anomalies. I. Pitch-intensity effects and diplacusis in normal ears,'' J. Acoust. Soc. Am. **72,** 1394–1402.

Burns, E. M., and Viemeister, N. F. (**1981**). ''Played again SAM: Further observations on the pitch of amplitude-modulated noise,'' J. Acoust. Soc. Am. **70,** 1655–1660.

Carlyon, R. P. (**1994**). ''Detecting pitch-pulse asynchronies and differences in fundamental frequency,'' J. Acoust. Soc. Am. **95,** 968–979.

Carlyon, R. P. (**1996**). ''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., 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.

Carlyon, R. P., Demany, L., and Semal, C. (**1992**). ''Detection of across-frequency differences in fundamental frequency,'' J. Acoust. Soc. Am. **91,** 279–292.

Cullen, J. K., and Long, G. R. (**1986**). ''Rate discrimination of high-pass filtered pulse trains,'' J. Acoust. Soc. Am. **79,** 114–119.

de Cheveigné, A. (**1993**). ''Separation of concurrent harmonic sounds: Fundamental frequency estimation and a time-domain cancellation model of auditory processing,'' J. Acoust. Soc. Am. **93,** 3271–3290.

Dobie, R. A., and Dillier, N. (**1985**). ''Some aspects of temporal coding for single-channel electrical stimulation of the cochlea,'' Hearing Res. **18,** 41–45.

Hall, J. W., Haggard, M. P., and Fernandes, M. A. (**1984**). ''Detection in noise by spectrotemporal pattern analysis,'' J. Acoust. Soc. Am. **76,** 50–56.

Hartmann, W. M. (**1988**). ''Pitch perception and the segregation and integration of auditory entities,'' in *Auditory Function*, edited by G. M. Edelman, W. E. Gall, and W. M. Cowan (Wiley, New York), pp. 623–645.

Hoekstra, A. (**1979**). ''Frequency discrimination and frequency analysis in hearing,'' Ph.D. Institute of Audiology, University Hospital, Groningen, Netherlands.

Houtsma, A. J. M., and Smurzynski, J. (**1990**). ''J. F. Schouten revisited: Pitch of complex tones having many high-order harmonics,'' J. Acoust. Soc. Am. **87,** 304–310.

Licklider, J. C. R. (**1951**). ''A duplex theory of pitch perception,'' Experientia **7,** 128–133.

McFadden, D. M. (**1987**). ''Comodulation detection differences using noise-band signals,'' J. Acoust. Soc. Am. **81,** 1519–1527.

Meddis, R., and Hewitt, M. (**1991**). ''Virtual pitch and phase sensitivity studied using a computer model of the auditory periphery: Pitch identification,'' J. Acoust. Soc. Am. **89,** 2866–2882.

Moore, B. C. J. (**1989**). *An Introduction to the Psychology of Hearing* (Academic, New York).

Patterson, R. D., Allerhand, M., 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.

Patterson, R. D., Milroy, R., and Allerhand, M. (**1993**). ''What is the octave of a harmonically rich note?,'' Contemp. Music Rev. **9**, 69–81.

Patterson, R. D., Nimmo-Smith, I., Holdsworth, J., and Rice, P. (**1988**). ''Spiral Vos Final Report, Part A: The Auditory Filterbank,'' Applied Psychology Unit, Cambridge, England, Contract Report: APU 2341.

Pijl, S., and Schwarz, D. W. F. (**1995**). ''Melody recognition and musical interval perception by deaf subjects stimulated with electrical pulse trains through single cochlear implant electrodes,'' J. Acoust. Soc. Am. **98,** 886–895.

Richards, V. M. (**1987**). ''Monaural envelope correlation perception,'' J. Acoust. Soc. Am. **82,** 1621–1630.

Ritsma, R. J. (**1963**). ''On pitch discrimination of residue tones,'' Int. Audiol. **2**, 34–37.

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.

Siebert, W. M. (**1968**). ''Stimulus transformations in the peripheral auditory system,'' in *Recognizing Patterns*, edited by P. A. Kolers and M. Eden (MIT, Cambridge, MA), pp. 104–133.

Slaney, M., and Lyon, R. F. (**1990**). ''A perceptual pitch detector,'' Proc. Int. Conf. Acoustics, Speech, and Signal Processing, pp. 357–360.

Zwicker, E. (**1952**). ''Die Grenzen der Hörbarkeit der Amplitudenmodulation und der Frequenzmodulation eines Tones,'' Acustica **4,** 415–420.

Modulation detection interference with two-component masker modulators

Stanley Sheft and William A. Yost
Parmly Hearing Institute, Loyola University of Chicago, 6525 N. Sheridan Road, Chicago, Illinois 60626

(Received 22 January 1997; accepted for publication 17 April 1997)

Thresholds were measured for detecting amplitude modulation of a single tonal carrier in the presence of modulated maskers. The masker modulator contained either one or two sinusoidal components. As found in previous studies, there was a detrimental effect of masker modulation on the detection of probe modulation. Interference was obtained with a 4-Hz separation between the two components of the masker modulator with the probe modulated at the 4-Hz beat rate. The amount of interference obtained with the two-component masker modulator was greater than that produced by the individual components of the masker modulator. The amount of interference increased with the depth of the masker-modulator beating and the number of masker carriers. The interference due to beating of the masker-modulator components was reduced when the beat rate was increased from 4 to 10 Hz. Results obtained with two-component masker modulators indicate a masking effect not predicted by a spectral representation of the probe and masker envelopes. [S0001-4966(97)01508-7]

PACS numbers: 43.66.Mk, 43.66.Fe [JWH]

INTRODUCTION

The ability to detect amplitude modulation (AM) of a tonal carrier can be disrupted by modulation in another spectral region (Yost and Sheft, 1989, 1994; Yost *et al.*, 1989; Hall and Grose, 1991; Bacon and Konrad, 1993; Bacon and Moore, 1993; Bacon and Opie, 1994). The elevation in probe modulation detection threshold due to the presence of a modulated masker has been termed modulation detection interference (MDI). Interference is also obtained when the task involves discrimination of probe modulation depth or rate (Yost *et al.*, 1989; Yost and Sheft, 1994; Moore *et al.*, 1991; Moore and Shailer, 1994). With sinusoidal amplitude modulation (SAM), MDI is generally, but not always, greatest when the probe and masker are modulated at the same rate (Yost *et al.*, 1989; Moore *et al.*, 1991; Bacon and Opie, 1994; Bacon *et al.*, 1995). The amount of interference then gradually declines with increasing separation between the probe and masker modulation rates. This broad tuning suggests that MDI in some way depends on the similarity between the probe and masker envelope fluctuations. When the probe and masker are modulated at the same rate, the envelope phase relationship has an inconsistent effect on MDI. Results from some studies show large individual differences in the effect of modulator phase (Yost and Sheft, 1989, Yost and Sheft, 1994; Bacon and Konrad, 1993) while other studies indicate only a small or absent phase effect (Moore *et al.*, 1991; Moore and Shailer, 1994). With narrow-band noise modulators, substantial MDI is present with both comodulated and independently modulated envelopes (Mendoza *et al.*, 1995; Moore *et al.*, 1995; Moore and Jorasz, 1996). These results indicate that envelope correlation is not a good metric for the modulation similarity underlying MDI.

The similarity between the probe and masker modulations can be approached in terms of either a temporal or spectral representation of the envelope fluctuations. If only the amplitude information from the envelope spectrum is processed and the phase information is discarded, the two representations become distinct. Studies of multitone masking and spectral shape discrimination have led to the suggestion that change in the power spectrum of the envelope can be used as a detection cue (Martens, 1982; Green *et al.*, 1992). Houtgast (1989) and Bacon and Grantham (1989) found that with wideband noise carriers, AM detection in the presence of competing modulation was poorest when the signal and masker modulation rates were the same with performance improving as the two rates diverged. These results led to speculation on the relevance of a spectral representation of complex envelope modulation. Recent modeling of auditory processing by Dau and his co-workers (Dau *et al.*, 1996) proposed analysis of the stimulus envelope by a modulation filterbank. The envelope spectrum is determined by this modulation-rate specific analysis.

Aspects of MDI are consistent with this modeling approach. To account for the tuning observed in the modulation domain, Yost *et al.* (1989) suggested that MDI may be a consequence of processing by modulation-rate specific channels. Yost and Sheft (1990) showed that a narrow-band noise masker centered at 1 kHz can significantly interfere with SAM detection for a 4-kHz pure-tone carrier. With a 32-Hz masker bandwidth, detection thresholds for SAM rates from 2 to 32 Hz were in the range of -7.5 to -5 dB in terms of $20 \log m$ where m is modulation depth. Without the masker, thresholds were in the range of -30 to -22.5 dB. The broad masking effect is consistent with the spectral representation of the masker envelope which extends from dc to frequencies above the noise bandwidth (Lawson and Uhlenbeck, 1950).

When the characteristics of the probe and masker modulation are reversed with modulation of the probe carrier by narrow-band noise and sinusoidal modulation of the masker, a different pattern of results is obtained. Moore *et al.* (1995) measured modulation depth discrimination thresholds for a sinusoidal carrier modulated by a 10-Hz-wide band of noise.

The masker was sinusoidally modulated at a rate ranging from 2 to 64 Hz. Results indicated a broadly tuned masking effect with the greatest amount of interference obtained when the masker AM rate was close to the mean modulation rate of the probe as determined by the expected number of envelope maxima per second. Since the spectrum of the probe envelope does not show a peak at the mean modulation rate, a spectral representation of the probe and masker envelopes would not predict the experimental results. A temporal representation based on the moment-to-moment fluctuations of the two envelopes would also not predict the results. Moore and his co-workers suggested that similarity in the long-term statistical properties of the probe and masker envelopes, such as the expected number of envelope maxima per second, determined the pattern of interference.

Shailer and Moore (1993) independently varied the envelope shape and periodicity of the probe and masker to determine the relative contribution of each to MDI. Envelope shape was varied with the modulator rise–fall time, the time for the cosinusoidal transition from the trough to peak values, and vice versa, of an envelope cycle. They found that when the probe and masker were modulated at the same rate of ten periods per second, MDI increased as the masker-modulator rise–fall time was shortened if the probe rise–fall time was constant. In other conditions, MDI increased as the probe-modulator rise–fall time increased when the masker rise–fall time was constant. MDI was therefore not greatest when the probe and masker modulations were most similar, regardless of whether similarity was defined in terms of temporal or spectral representations of the two envelopes. MDI can clearly be influenced by factors other than the similarity between the probe and masker envelopes. Shailer and Moore observed that the pattern of results is consistent with the notion that the carrier with the shorter rise–fall time is perceptually dominant. Since variation in the envelope rise–fall time is reflected in the envelope spectrum, broadening as rise–fall time is shortened, the effect of rise–fall time can be considered in terms of either a temporal or spectral representation of the envelope.

The intent of the present study was to measure MDI with a stimulus configuration in which temporal and spectral representations of the probe and masker envelopes would predict differing amounts of interference. Masker modulators contained two sinusoidal components. With these modulators, the masker envelope exhibits a periodicity at the beat rate of the modulator that is not represented in its envelope spectrum. By assessing the ability to detect probe modulation at the beat rate of the masker modulator, results are used to evaluate the extent to which the temporal pattern or spectral composition of the envelope is the primary determinant of the interference effects.

I. METHOD

A. Stimuli

Thresholds were measured for detecting amplitude modulation of a probe tone in the presence of modulated masker tones. The masker modulator contained either one or two sinusoidal components. The stimuli can be represented as the sum of the probe and masker complexes such that

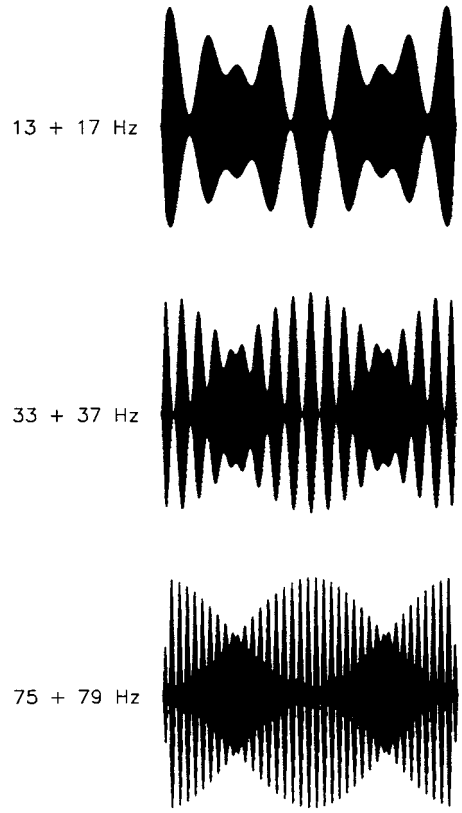


FIG. 1. A schematic diagram showing three examples of a 500-ms masker carrier modulated by a two-component complex. The frequencies of the two sinusoids of the modulator are indicated to the left of each waveform. The amplitude of each component is equal to half the dc level. In all three examples, the modulator components beat at a 4-Hz rate. The carrier fine structure is not visible in any of the schematic representations.

$$s(t)=A_p[1+m_p \sin(2\pi f_p t)]c_p(t)+A_m[1 + m_{m1} \sin(2\pi f_{m1} t)+m_{m2} \sin(2\pi f_{m2} t)]c_m(t), \quad (1)$$

where m_p, f_p, and $c_p(t)$ are the probe modulation index, rate, and carrier, respectively, m_{m1} and f_{m1} are the modulation index and rate of one component of the masker modulator and m_{m2} and f_{m2} are the modulation index and rate of the other component, and $c_m(t)$ is the masker carrier. A_p and A_m are scaling factors that were varied to maintain constant average power of $s(t)$ across changes in modulation depth (see Viemeister, 1979). When the masker modulation was sinusoidal, m_{m2} and f_{m2} were equal to zero. In conditions with two sinusoidal components in the masker modulator, the sum of the m_{m1} and m_{m2} never exceeded 1.0 to avoid overmodulation. Examples of modulated-masker waveforms are shown in Fig 1.

The probe modulation rate f_p was either 4 or 10 Hz. When f_p was equal to the beat rate of a masker modulator ($f_p=f_{m2}-f_{m1}$), the phase of the masker envelope beat cycle led the probe envelope phase by 90° (i.e., for the waveforms shown in Fig. 1 with the beat cycle beginning at a peak, the probe modulator would be in sine phase). The

probe carrier frequency was always 909 Hz. The masker carrier $c_m(t)$ consisted of one to three equal-amplitude sinusoids. Possible masker-carrier frequencies were 1414, 2432, and 4186 Hz. The level of each component of the probe and masker carriers was fixed at 70 dB SPL. Stimuli were 500 ms in duration and shaped by 10-ms $\cos^2$ rise–fall ramps.

A Masscomp minicomputer interfaced with a 16-bit digital-to-analog converter was used for stimulus generation and experimental control. Stimuli were generated at a 20-kHz sampling rate. Following analog conversion, all stimuli were low-pass filtered (Krohn-Hite, model 3343R) at 8 kHz. Stimuli were presented monaurally to the left ear through a TDH-49 earphone with the subjects seated in a double-walled soundproof booth.

B. Procedure

Thresholds were measured using a two-interval two-alternative forced-choice (2IFC) tracking procedure (Levitt, 1971). The two observation intervals were separated by 500 ms. Feedback was provided after each trial. The task was to identify the interval containing the probe modulation f_p. The modulation index (m_p) was decreased by 2 dB in steps of 20 log m_p following two consecutive correct responses and increased by 2 dB following one incorrect response. A run was terminated after 50 trials. Thresholds are reported in terms of 20 log m, where m equals m_p. Runs containing less than ten reversals were discarded. Excluding the first four reversals (five if the total number of reversals was odd), threshold was estimated as the average of the levels in dB of the remaining reversals. Reported threshold values are the mean of at least six threshold estimates per condition.

C. Subjects

Data were collected from seven subjects with pure-tone thresholds within 10 dB of ANSI (1989) standard audiometric thresholds. Subjects received a minimum of 12 h of practice before data collection began. Three to four subjects participated in each experimental condition.

II. RESULTS

A. Effects of masker-modulator beat rate and depth

If MDI results from the probe and masker sharing a similar envelope periodicity, the two-component masker modulator could interfere with probe AM detection at the masker-modulator beat rate. If the similarity in modulation underlying MDI is described by spectral representations of the probe and masker envelopes, interference is not predicted except to the extent that it is produced by the individual masker-modulator components. The initial conditions evaluated probe AM detection at the masker-modulator beat rate. The depth of the masker envelope fluctuation at the beat rate was controlled by varying the amplitude (m_{m2}) of one of the modulator components while the amplitude of the other component (m_{m1}) was fixed at 0.5. The frequency of the fixed-amplitude component was 33 Hz. The frequency of the variable-amplitude component was either 37 or 43 Hz. The probe modulation rate was equal to the beat rate of the masker modulator, either 4 or 10 Hz. A three-tone complex of frequencies 1414, 2432, and 4186 Hz was used as the masker carrier.

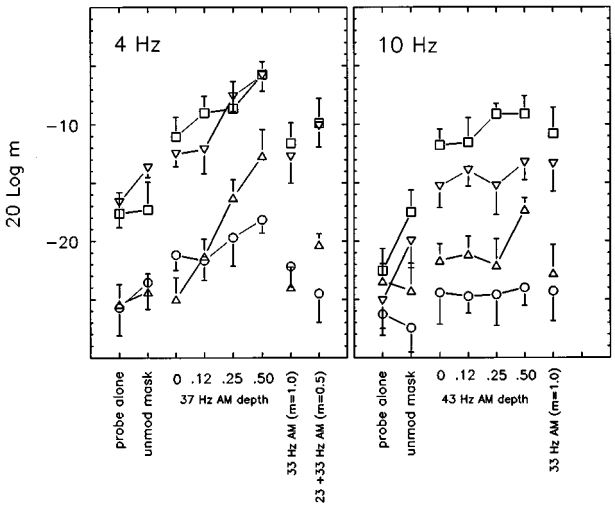


FIG. 2. Thresholds for detecting either 4-Hz (left panel) or 10-Hz (right panel) probe SAM. The probe carrier frequency was 909 Hz and the masker carrier consisted of three tones (1414+2432+4186 Hz). Different configurations for the masker modulator are marked along the abscissa. Symbol type indicates subject number: S1 (triangles), S2 (circles), S3 (squares), S4 (inverted triangles). Error bars represent one standard deviation from the mean threshold.

Thresholds for detecting either 4- or 10-Hz SAM of the 909-Hz probe carrier are shown in Fig. 2. Different symbol shapes are used for the results from each of the four subjects. On the left side of each panel are the thresholds obtained when no masker was present and when the masker was not modulated. Looking first at the results for detecting 10-Hz probe modulation (right panel), thresholds without the masker present ranged from roughly −26.5 to −22.5 dB. Addition of the unmodulated masker raised thresholds by more than 5 dB for subjects S3 (squares) and S4 (inverted triangles) while it had little effect on subjects S1 (triangles) and S2 (circles). Several past studies have shown that the addition of an unmodulated masker can interfere with the detection of probe modulation with individual differences evident in the results (Bacon and Moore, 1993; Bacon and Opie, 1994). Subjects S3 and S4 also exhibited greater increases in threshold than the other two subjects when the masker was sinusoidally modulated at 33 Hz. The presence of MDI in these conditions with a relatively wide separation between the probe and masker AM rates is consistent with past work showing relatively broad selectivity to masker AM rate when measuring interference (Yost *et al.*, 1989; Moore *et al.*, 1991; Bacon and Opie, 1994). For all subjects, thresholds were roughly the same when the modulation index of the 33-Hz sinusoidal modulator was either 0.5 or 1.0. Bacon and Konrad (1993) found a near linear relationship between logarithmic measures of masker-modulation index and probe AM detection threshold when the probe and masker were both modulated at 10 Hz. The present results suggest that this relationship may not hold when the probe and masker are modulated at different rates.

The key conditions shown in the right panel of Fig. 2

involved modulation of the masker by both the 33- and 43-Hz components. Except for subject S1 (triangles), thresholds were relatively stable across change from 0.0 to 0.5 in the amplitude m_{m2} of the 43-Hz modulator in the four conditions in which the amplitude m_{m1} of the 33-Hz modulator was constant at 0.5. A repeated measures analysis of variance (ANOVA)was performed on factor m_{m2}. The effect of m_{m2} was not statistically significant at the 0.05 level. This result indicates that the presence of masker modulator beating at the 10-Hz modulation rate of the probe does not affect probe AM detection ability.

A different pattern of results was obtained in the conditions measuring detection of 4-Hz probe modulation, as shown in the left panel of Fig. 2. The unmodulated masker and also the masker sinusoidally modulated at 33 Hz had a smaller effect on 4- than 10-Hz AM detection thresholds. More importantly, the 4-Hz beating of the masker modulators at the 4-Hz rate of probe modulation led to an increase in MDI. As amplitude m_{m2} of the 37-Hz modulator increased from 0.0 to 0.5 with the amplitude m_{m1} of the 33-Hz modulator constant at 0.5, thresholds for detecting 4-Hz SAM of the probe carrier rose for all four subjects. A repeated measure ANOVA confirmed that the effect of m_{m2} was significant ($F_{3,9}=10.8$, $p=0.002$). The additional interference is not directly related to the increase in masker modulator power as m_{m2} increases. A sinusoidal modulator with an amplitude of 1.0 leads to 3-dB greater modulator power than a two-component modulator with the amplitude of each component at 0.5. When the masker was sinusoidally modulated at 33 Hz with a modulation index of 1.0, thresholds were lower than those obtained with any of the two-component modulators.

The interference obtained with the two-component masker modulator was also not simply due to the presence of a masker-envelope beat fluctuation regardless of rate. In one condition, the frequency f_{m2} of the second masker modulator was changed from 37 to 23 Hz so that the modulator would no longer beat at the probe modulation rate. The amplitude of each component of the masker modulator was 0.5. Thresholds from this condition are shown on the far right of the left panel of Fig. 2. Though one component of the masker modulator was now closer to the 4-Hz modulation rate of the probe, the amount of interference decreased when the masker-modulator beat rate was increased from 4 to 10 Hz, indicating that the interference was linked to beating at the modulation rate of the probe.

B. Effect of the number of masker carriers

In the first set of conditions, a three-tone complex was used as the masker carrier with all the masker carrier frequencies above the probe carrier frequency. Results from past studies that have used sinusoidal modulators indicate that MDI is not influenced by the number of masker carriers as long as the proximity of the masker carrier(s) closest to the probe carrier frequency does not vary (Bacon and Moore, 1993; Moore and Shailer, 1994; Bacon *et al.*, 1995). A recent study by Moore and Jorasz (1996) measured MDI as a function of the number of masker carriers with narrow-band noise modulators. Across conditions in which the proximity

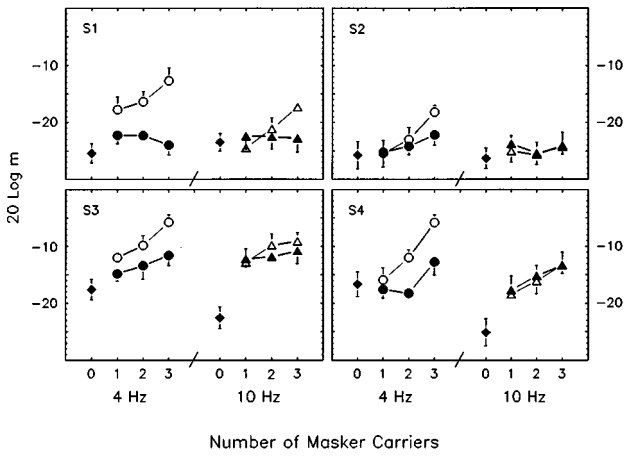


FIG. 3. Thresholds for detecting either 4- or 10-Hz SAM of a 909-Hz probe carrier. The abscissa indicates the number of masker carriers at the following frequencies: 0 (probe alone), 1 (1414 Hz), 2 (1414+2432 Hz), 3 (1414 +2432+4186 Hz). The parameter is masker condition: no masker (filled diamonds), 33-Hz AM (filled circles and filled triangles), 33+37 Hz AM (open circles), 33+43 Hz AM (open triangles). Error bars represent one standard deviation from the mean threshold.

of the closest masker carrier to probe carrier frequency was constant, an effect of the number of masker carriers was obtained in some cases. The difference in the effect of the number of masker carriers among these studies may in part relate to whether the modulators were sinusoidal or complex.

In the second set of conditions, the number of masker carriers was varied in conditions which used both sinusoidal and two-component masker modulators. Carriers were added to the masker progressively farther from the 909-Hz probe carrier frequency. The masker had either one (1414 Hz), two (1414 and 2432 Hz), or three (1414, 2432, and 4186 Hz) carriers. The masker modulator was either a 33-Hz sinusoid or a two-component complex. The frequencies of the two components of the masker modulator were 33 and 37 Hz, or 33 and 43 Hz. When modulated by the 33-Hz sinusoid, the masker modulation index was 1.0. With the two-component modulator, the amplitude of each component (m_{m1} and m_{m2}) was 0.5.

AM detection thresholds for the 909-Hz probe carrier as a function of the number of masker carriers are shown in Fig. 3. Individual results are shown in separate panels. Results obtained with either none or three masker carriers are replotted from Fig. 2. The left side of each panel shows thresholds for 4-Hz probe modulation and the right side for 10 Hz. The filled diamonds indicate thresholds for detecting either 4- or 10-Hz probe modulation without masking. Thresholds for detecting 4- and 10-Hz modulation in the presence of 33-Hz masker modulation are shown with the filled circles and filled triangles, respectively. The open symbols show the thresholds from conditions in which the two-component masker modulator was beating at the probe modulation rate.

The number of masker carriers had relatively little effect when the masker was sinusoidally modulated at 33 Hz (filled symbols). In these conditions, the largest change in threshold as a function of the number of masker carriers was roughly 5 dB for subject S4. There was also only a small effect of the number of masker carriers on the detection of 10-Hz

probe modulation. Consequently, similar thresholds were obtained in the presence of the one- and two-component masker modulators with the 10-Hz rate of probe modulation. This is indicated by the overlap of the functions marked with the open and filled triangles. The one exception to this trend was for subject S1 with three masker carriers; threshold in the presence of the two-component masker modulator was roughly 5-dB higher than that obtained with sinusoidal masker modulation. This was also the only threshold from the data of Fig. 2 that indicated additional interference with the masker-modulator beating at the 10-Hz rate of probe modulation.

The effect of the number of masker carriers was largest with the two-component masker modulator beating at the 4-Hz probe modulation rate. The functions marked with the open circles have the steepest slopes. The amount of additional interference due to masker-modulator beating is indicated by the separation between the open and filled symbols of the same shape. With 4-Hz probe modulation, interference due to masker-modulator beating increased with the number of masker carriers. Subjects S2 and S4 showed interference due to modulator beating only with multiple-masker carriers. For these two subjects, the functions marked with the open and filled circles do not diverge until there were two to three masker carriers.

Two separate repeated measures ANOVAs were conducted with factors masker-modulator type (sinusoidal or two component) and number of masker carriers. One ANOVA was for the thresholds obtained with the 4-Hz rate of probe modulation and the other was for the 10-Hz results. For modulation detection at 10 Hz, the main effect of modulator type was not significant while the effect of the number of masker carriers was significant ($F_{2,6}=7.08$, $p=0.026$). *Post-hoc* pairwise comparisons indicated that the only significant difference ($p<0.05$) was between the two-component-modulator conditions in which there were one or three masker carriers.

For the detection of 4-Hz modulation, the main effect of masker-modulator type was significant ($F_{1,3}=15.34$, $p=0.03$). Thresholds were significantly higher with the two-component rather than sinusoidal masker modulator indicating additional interference due to masker-modulator beating. There was also a significant main effect of number of masker carriers ($F_{2,6}=14.43$, $p=0.005$). *Post-hoc* pairwise comparisons indicated that the effect of the number of masker carriers was restricted to the two-component-modulator conditions with all three comparisons (1 vs 2, 2 vs 3, and 1 vs 3 masker carriers) showing a significant difference ($p<0.05$). Results obtained at both modulation rates are then in agreement with previous studies showing no effect on MDI of adding progressively distal masker carriers when using sinusoidal modulators (Bacon and Moore, 1993; Moore and Shailer, 1994; Bacon *et al.*, 1995). For 4-Hz modulation detection, the interaction between factors masker-modulator type and number of masker carriers was significant ($F_{2,6}=16.22$, $p=0.004$), reflecting the increase in interference due to modulator beating with number of masker carriers. Thus in the present study as in the work of Moore and Jorasz (1996), an effect of the number of masker carriers was obtained only when using a complex modulator.

TABLE I. Thresholds in dB for detecting 4-Hz modulation of a 909-Hz carrier with standard deviations listed in parenthesis below each threshold. When present, the masker complex consisted of three carriers, 1414, 2432, and 4186 Hz. The masker modulation index was 1.0 when the modulator was sinusoidal (e.g., 15 Hz). When the masker modulator was a two-component complex (e.g., 13+17 Hz), the amplitude of each component was 0.5.

			Masker-modulator frequency (Hz)			
Subject	No mask	Unmodulated mask	15	13+17	77	75+79
S5	−23.3 (1.9)	−21.8 (3.3)	−9.6 (2.9)	−2.9 (0.9)	−15.0 (2.7)	−12.7 (3.0)
S6	−25.8 (1.9)	−17.7 (1.8)	−8.9 (1.6)	−5.0 (1.1)	−17.3 (2.7)	−14.1 (3.5)
S7	−28.0 (1.6)	−22.5 (2.0)	−15.3 (1.6)	−8.6 (1.5)	−23.4 (2.4)	−13.1 (2.8)

C. Effect of masker-modulator component frequencies

The final set of AM detection conditions evaluated the effect of the frequencies of the modulator components with the beat rate constant at 4 Hz. The components f_{m1} and f_{m2} of the masker modulator were either 13 and 17 Hz, or 75 and 79 Hz. The amplitude of each component (m_{m1} and m_{m2}) was 0.5. Thresholds were also measured with sinusoidal masker modulation at 15 Hz alone and 77 Hz alone with a modulation index of 1.0. The three-tone complex of frequencies 1414, 2432, and 4186 Hz was used as the masker carrier.

Thresholds for detecting 4-Hz probe modulation are shown in Table I for three new subjects. For these subjects, thresholds without the masker and in the presence of the unmodulated masker fell within the range exhibited by subjects S1–S4 in the initial set of conditions (see Fig. 2). Masker modulation at 15 Hz elevated thresholds for all three subjects. When referenced to the unmodulated-masker thresholds, 77-Hz masker modulation led to MDI for only subject S5. All subjects exhibited MDI for both two-component modulators. The threshold shifts due to masker modulation were greater with the two-component modulator than with the proximal sinusoidal modulator. The increased interference occurred despite the 3-dB drop in masker-modulator power with the two-component modulator.

The data of Fig. 2 and Table I show for the two separate groups of subjects 4-Hz AM detection thresholds obtained in the presence of three different two-component modulators with a 4-Hz beat rate. For these modulator pairs, thresholds were poorest when the frequencies of the modulator components were 13 and 17 Hz. However, 15-Hz SAM of the masker also resulted in substantial MDI. Subtracting thresholds obtained with sinusoidal modulation (m_{m1} equal to 1.0) of the masker from those obtained with the two-component modulator (m_{m1} and m_{m2} equal to 0.5) gives one measure of the additional interference due to the envelope beat fluctuation. This metric eliminates the effect on MDI of the proximity of the masker-modulator component frequencies to the modulation rate of the probe. The average additional inter-

ference was 5.8 dB with the 13- and 17-Hz modulator pair, 7 dB with the 33- and 37-Hz pair, and 5.2 dB with the 75- and 79-Hz pair. With this metric, the three modulator pairs resulted in roughly the same amount of additional interference due to modulator beating. Though the three masker modulators all have a 4-Hz beat periodicity, the three masker envelopes are distinguished by the rate of their more rapid envelope fluctuation (see Fig. 1). The rate of this more rapid envelope fluctuation increases with the component frequencies of the masker modulator. The roughly constant amount of additional interference for the three modulators suggests that the common periodicity between the probe modulator and the masker beating rather than the ''fine structure'' of the temporal pattern of the masker envelope determines the amount of additional MDI due to masker-modulator beating.

III. DISCUSSION

The purpose of the present study was to evaluate envelope characteristics that might be responsible for MDI. Previous studies reviewed in the Introduction indicate that the interference is often dependent on some aspect of the similarity between the probe and masker envelope fluctuations. The question then is how to quantify the similarity in a manner consistent with the amount of interference in various conditions. Two-component masker modulators were used in the present work in an attempt to distinguish between the envelope spectrum and the temporal pattern of envelope fluctuation as the basis of MDI. The primary finding of this study was that interference due to beating of the masker-modulator components was obtained with a 4-Hz separation between modulator components. This interference was not present when the beat rate was increased to 10 Hz.

Amplitude modulation of a carrier by a complex that contains two sinusoidal components results in an envelope periodicity at the beat rate of the modulator components that is not represented in the envelope spectrum. Psychophysical models of envelope detection utilize half-wave rectification followed by low-pass filtering (e.g, Viemeister, 1979). As long as the carrier is not overmodulated, half-wave rectification does not introduce distortion of the modulating signal (Russell, 1962). Half-wave rectification followed by low-pass filtering thus functions as a linear envelope detector. To generate in the detector output a spectral component at the beat rate of a two-component modulator, a second processing nonlinearity is needed. Shofner *et al.* (1996) recorded the responses of neurons in the chinchilla cochlear nucleus to best-frequency tones modulated by two-component complexes. Fourier analysis of the post-stimulus time histograms showed spectral peaks corresponding to the envelope beat frequencies. This distortion of the modulating signal was attributed to the compressive nonlinearity of the neural input-output function. For a fixed frequency separation between modulator components, the effect of this neural nonlinearity does not vary with component frequencies. With a 4-Hz separation between modulator components in the present study, the amount of MDI due to masker-modulator beating did not vary markedly with modulator component frequencies (see Sec. II C).

In the physiological study of Shofner *et al.* (1996), the magnitude of the spectral peak at the beat frequency was independent of beat rate for rates from 4 to 16 Hz. Thus while demonstrating a possible basis for the psychophysically measured MDI at the masker-modulator beat rate, the physiology does not account for the decrease in the amount of MDI at the beat rate when the rate was increased from 4 to 10 Hz. The same difficulty exists with using the similarity of pattern of envelope fluctuation as a basis for the interference; this would not predict an effect of beat rate in the range of 4 to 10 Hz. Yost *et al.* (1989) and Bacon and Konrad (1993) measured MDI with sinusoidal modulators as a function of the common modulation rate of the probe and masker. Results from both studies showed that the amount of interference decreased as the modulation rate was increased from 5 to 10 Hz. The effect of masker-modulator beat rate in the present study may then in part reflect a general effect that is not unique to beating modulators. It is worth noting, however, that while the amount of MDI obtained with sinusoidal modulators in the previous studies decreased as modulation rate was raised, interference was still present with 10-Hz modulation of the probe and masker. In the present study, the effect of masker-modulator beating was not significant at the 10-Hz beat rate.

Interference was obtained at the beat rate of the masker modulator only with multiple-masker carriers with less than an octave separation between the probe carrier and the nearest masker carrier. Bacon and Moore (1993) showed that a within-channel interaction between the probe and masker modulations is likely with a masker carrier less than an octave above the probe carrier frequency. With the 4-Hz masker modulator beat rate, adding progressively distal masker carriers led to greater interference in detecting 4-Hz probe modulation. This result suggests involvement of across-channel modulation processing. The intent of the present study was to evaluate envelope characteristics that may underlie MDI. An assumption of this work was that the envelope characteristics potentially underlying MDI could be studied in conditions in which both within- and across-channel interactions of modulation may occur (see Bacon and Konrad, 1993). While the stimulus configuration allowed for a distinction between spectral and temporal representations of the complex masker envelope, the effect of masker-modulator beat rate did not distinguish either as the primary determinant of MDI.

IV. CONCLUSIONS

This study examined AM detection in the presence of two-component masker modulators. Interference was obtained when the probe was modulated at the beat frequency of the masker modulator. The amount of interference due to beating of the masker modulator was reduced when the beat rate was increased from 4 to 10 Hz. At the 4-Hz beat rate, the amount of interference increased with the depth of masker-modulator beating and the number of masker carriers. With a 4-Hz separation between modulator components, the amount of additional interference due to modulator beating did not

depend markedly on the modulator component frequencies. The reason why the effects were observed only at 4 Hz and not at 10 Hz is not known.

ACKNOWLEDGMENTS

This research was supported by a program project grant from the National Institutes of Health. We would like to thank our colleagues at the Parmly Hearing Institute for comments on our work, and also Brian Moore, Sid Bacon, and Jungmee Lee for comments on an earlier version of the manuscript.

ANSI (**1989**). ANSI S3.6-1989, ''Specifications for audiometers'' (American National Standards Institute, New York).

Bacon, S. P., and Grantham, D. W. (**1989**). ''Modulation masking: Effects of modulation frequency, depth, and phase,'' J. Acoust. Soc. Am. **85,** 2575–2580.

Bacon, S. P., and Konrad, D. L. (**1993**). ''Modulation detection interference under conditions favoring within- or across-channel processing,'' J. Acoust. Soc. Am. **93,** 1012–1022.

Bacon, S. P., and Moore, B. C. J. (**1993**). ''Modulation detection interference: Some spectral effects,'' J. Acoust. Soc. Am. **93,** 3442–3453.

Bacon, S. P., and Opie, J. M. (**1994**). ''Monotic and dichotic modulation detection interference in practiced and unpracticed subjects,'' J. Acoust. Soc. Am. **95,** 2637–2641.

Bacon, S. P., Moore, B. C. J., Shailer, M. J., and Jorasz, U. (**1995**). ''Effects of combining maskers in modulation detection interference,'' J. Acoust. Soc. Am. **97,** 1847–1853.

Dau, T., Kollmeier, B., and Kohlrausch, A. (**1996**). ''A quantitative prediction of modulation masking with an optimal-detector model,'' J. Acoust. Soc. Am. **99,** 2565.

Grantham, D. W., and Bacon, S. P. (**1991**). ''Binaural modulation masking,'' J. Acoust. Soc. Am. **89,** 1340–1349.

Green, D. M., Berg, B. G., Dai, H., Eddins, D. A., Onsan, Z., and Nguyen, Q. (**1992**). ''Spectral shape discrimination of narrow-band sounds,'' J. Acoust. Soc. Am. **92,** 2586–2597.

Hall, J. W., and Grose, J. H. (**1991**). ''Some effects of auditory grouping factors on modulation detection interference (MDI),'' J. Acoust. Soc. Am. **90,** 3028–3035.

Houtgast, T. (**1989**). ''Frequency selectivity in amplitude-modulation detection,'' J. Acoust. Soc. Am. **85,** 1676–1680.

Lawson, J. L., and Uhlenbeck, G. E. (**1950**). *Threshold Signals* (McGraw-Hill, New York).

Levitt, H. (**1971**). ''Transformed up-down methods in psychoacoustics,'' J. Acoust. Soc. Am. **49,** 467–477.

Martens, J. P. (**1982**). ''A new theory for multitone masking,'' J. Acoust. Soc. Am. **72,** 397–405.

Mendoza, L., Hall, J. W., and Grose, J. H. (**1995**). ''Modulation detection interference using random and sinusoidal amplitude modulation,'' J. Acoust. Soc. Am. **97,** 2487–2492.

Moore, B. C. J., and Jorasz, U. (**1996**). ''Modulation discrimination interference and comodulation masking release as a function of the number and spectral placement of narrow-band modulators,'' J. Acoust. Soc. Am. **100,** 2373–2381.

Moore, B. C. J., and Shailer, M. J. (**1994**). ''Effects of harmonicity, modulator phase, and number of masker components on modulation discrimination interference,'' J. Acoust. Soc. Am. **95,** 3555–3560.

Moore, B. C. J., Sek, A., and Shailer, M. J. (**1995**). ''Modulation discrimination interference for narrow-band noise modulators,'' J. Acoust. Soc. Am. **97,** 2493–2497.

Moore, B. C. J., Glasberg, B. R., Gaunt, T., and Child, T. (**1991**). ''Across-channel masking of changes in modulation depth for amplitude- and frequency-modulated signals,'' Quart. J. Exp. Psych. **43A**, 327–347.

Russell, G. M. (**1962**). *Modulation and Coding in Information Systems* (Prentice-Hall, Englewood Cliffs, NJ).

Shailer, M. J., and Moore, B. C. J. (**1993**). ''Effects of modulation rate and rate of envelope change on modulation discrimination interference,'' J. Acoust. Soc. Am. **94,** 3138–3143.

Shofner, W. P., Sheft, S., and Guzman, S. J. (**1996**). ''Responses of ventral cochlear nucleus units in the chinchilla to amplitude modulation by low-frequency, two-tone complexes,'' J. Acoust. Soc. Am. **99,** 3592–3605.

Viemeister, N. F. (**1979**). ''Temporal modulation transfer functions based upon modulation thresholds,'' J. Acoust. Soc. Am. **66,** 1364–1380.

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., and Sheft, S. (**1990**). ''Modulation detection interference with narrow-band noise probes and maskers,'' paper presented at the Association for Research in Otolaryngology, St. Petersburg(unpublished).

Yost, W. A., and Sheft, S. (**1994**). ''Modulation detection interference: Across-frequency processing and auditory grouping,'' Hearing Res. **79,** 48–58.

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.

The effects of randomizing values of interaural disparities on binaural detection and on discrimination of interaural correlation

Leslie R. Bernstein and Constantine Trahiotis
Surgical Research Center, Department of Surgery (Otolaryngology) and Center for Neurological Sciences, University of Connecticut Health Center, Farmington, Connecticut 06030

(Received 7 January 1996; revised 28 April 1997; accepted 6 May 1997)

The purpose of this study was to assess whether randomizing (roving) baseline interaural parameters of binaural stimuli would adversely affect performance in masking-level difference and in interaural correlation-discrimination paradigms. Listeners' detection thresholds were measured in NoSo and NoSπ configurations for both broadband (100–3000 Hz) and narrow-band (450–550 Hz) maskers centered at 500 Hz. In addition, thresholds of interaural decorrelation (from a reference correlation of 1.0) were measured for 100-Hz-wide bands of noise centered at 500 Hz. Data were obtained under conditions in which either values of ITD or values of IID were roved both within and across trials. Data were also collected in the absence of a rove. The range of the rove was ± 500 μs for ITDs and was ± 6 dB for IIDs. The duration of the stimuli was varied between 17 and 310 ms. Overall, the results indicate that, independent of duration, roving the interaural cues produced small degradations of performance as compared to data obtained in the absence of a rove. The data are inconsistent with the notion that binaural detection depends upon reliable changes in laterality produced by adding Sπ signals to diotic maskers. On the other hand, the data are consistent with modern models of binaural processing. © *1997 Acoustical Society of America.*
[S0001-4966(97)06308-X]

PACS numbers: 43.66.Pn [JWH]

INTRODUCTION

A ubiquitous characteristic of experiments concerning the binaural release from masking is that the addition of the signal to the masker produces changes in interaural information that serve as cues for detection. It is typically the case that the interaural parameters of the masker are held constant both within and across trials. From a perceptual point of view, presentation of the masker typically results in an intracranial image having a position (determined by its interaural parameters) that is stable both within and across trials. The stability of the perceived locus of the masker can foster its use a ''reference'' against which the perceptual consequences of presentation of the signal-plus-masker can be compared (e.g., McFadden, 1966). In perceptual terms, one can conceive of the addition of the signal to the masker as being detected on the basis of a change in the laterality of the intracranial image in the spectral region (i.e., critical band) of the masker corresponding to the frequency of the signal. In fact, Hafter *et al.* (1969) demonstrated that listeners' ability to detect the addition of a 500-Hz, antiphasic (Sπ) tone to a diotic (No) broadband masker could be predicted from the (separately obtained) listeners' judgments concerning the perceived laterality of the pitch of the signal. Hafter *et al.*'s (1969) sophisticated analysis of their data indicates that, under certain conditions, changes in perceived laterality produced by adding the signal to the masker can be *sufficient* to account for its detectability.

This does not mean, however, that reliable and valid changes in perceived laterality are theoretically or logically *necessary* for the detection of the signal in the NoSπ configuration or, for that matter, in any of the other stimulus conditions that produce a release from masking. For example, Hafter *et al.* (1969) themselves discuss how Durlach's (1963) equalization-cancellation (EC) model can account for binaural detection without any reference whatsoever to intracranial space or location. Furthermore, the addition of the signal to the masker produces dynamically changing interaural temporal differences (ITDs) and dynamically changing interaural intensitive differences (IIDs) and the listeners' use of such cues may account for the detectability of the signal without requiring recourse to changes in lateral position, *per se*. Both Henning (1973) and Zurek (1991) have provided quantitative information concerning the distributions of ITD and IID that result from the addition of the signal to the masker.

The purpose of this study was to examine empirically whether and to what degree binaural detection would be affected under conditions in which the addition of the signal *could not* produce an informative change in laterality. This was accomplished by altering the ''standard'' paradigm such that a randomly chosen value of ITD or IID was imposed on both masker and signal-plus-masker waveforms, both within and across trials. Roving the baseline interaural cues in this manner precludes the use of the intracranial locus of the masker as a ''reference'' and vitiates the use of changes in lateral position, *per se*, as cues for detection of the signal. At the same time, changes in the widths of the distributions of the interaural cues produced by adding the signal are unaffected by roving the waveforms in this manner. By comparing binaural detection thresholds obtained under roving and fixed conditions, our experimental paradigm provided an op-

portunity to determine whether informative changes in laterality, *per se*, produced by the addition of the signal to the masker are necessary as well as sufficient.

I. EXPERIMENT 1

A. Procedure

Listeners' detection thresholds for 500-Hz tones were measured utilizing the interaural configurations NoSo and NoSπ. Both broadband (100–3000 Hz) and narrow-band (100-Hz-wide, centered at 500 Hz) noises served as maskers. When the broadband masker was employed, three combinations of masker/signal duration (including 10-ms $\cos^2$ rise/decay ramps) were used: 35 ms/25 ms, 145 ms/125 ms, and 310 ms/300 ms. When the narrow-band masker was employed, one additional combination of masker/signal duration was employed: 17 ms/12 ms (including 5-ms $\cos^2$ rise/decay ramps). Signals were always temporally centered within the maskers and were gated on and off within the plateau of the gating function of the masker.

All stimuli were generated digitally with a sampling rate of 20 kHz via a TDT array processor. The noises were constructed using inverse FFTs that yielded 32 768 points in the time domain. Tonal signals were generated in the time domain and added to noise waveforms prior to D/A conversion. All stimuli were low-pass filtered at 8.5 kHz (TDT FLT2) and presented via TDH-39 earphones (mounted in MX/41-AR cushions) to listeners seated in single-walled, IAC booths. Three young adults, two males and one female, with no evidence or history of hearing loss served as listeners. All listeners received extensive practice before collection of data began.

The stimuli were presented in a two-alternative temporal forced-choice adaptive task. Each trial consisted of a warning interval (500 ms) and two observation intervals separated by 450 ms. The observation intervals were marked by a visual display on a computer monitor. The tonal signal was presented with equal *a priori* probability in either the first or second interval. Independent masker waveforms were presented during each and every observation interval.

The level of the signal was varied adaptively in order to estimate 70.7% correct (Levitt, 1971). The initial step size for the adaptive track was 2 dB and was reduced to 1 dB after two reversals. A run was terminated after 12 reversals and threshold was defined as the average level of the signal across the last ten reversals.

All conditions of the experiment (interaural configuration×bandwidth×duration) were run in both the presence and the absence of roving interaural disparities. When ITD was roved, a randomly chosen value of ITD was imposed on both masker and signal-plus-masker waveforms, with the magnitude of the ITD being independently selected for each interval of each trial. The value of the ITD was chosen randomly from a range of ± 500 μs in 50-μs steps. A particular ITD was implemented by appropriately choosing ''later'' digital samples for the waveform destined to the delayed ear. Delays were imposed prior to applying rise/decay ramps. That is, the delays were ''ongoing'' and there were no onset/offset delays. When IID was roved, the value of IID was chosen randomly from a range of ± 6 dB in 0.5-dB steps and the IID was implemented via TDT programmable attenuators by an increase in the level of the stimulus by IID/2 in one ear and a decrease of IID/2 in the other. The spectrum level of the masking noise before any IID was imposed was equivalent to 50 dB SPL. Roving the ITDs and IIDs as described produced interval-by-interval variation of the lateral position of the stimuli, which were heard in one of many positions from far toward the left ear to far toward the right ear.

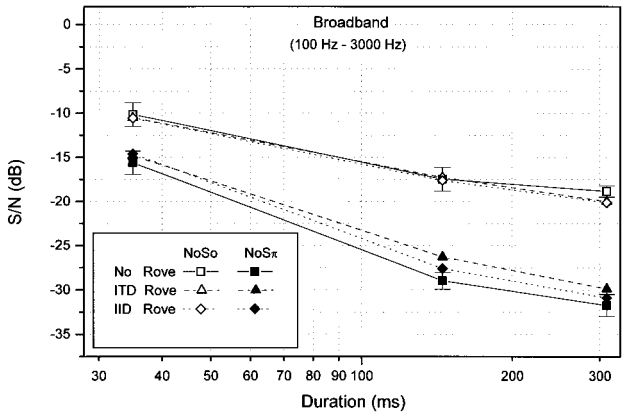


FIG. 1. Threshold S/N (in dB) as a function of masker duration obtained with the broadband (100–3000 Hz) masker. Open and closed symbols represent data obtained with the NoSo and NoSπ configurations, respectively, under conditions of no-rove (squares), a ± 500-μs ITD rove (triangles), and a ± 6-dB IID rove (diamonds). The data points represent the average across three listeners. Error bars represent ± 1 standard error of the mean.

Experimental conditions were assigned in quasi-random order and, during each 1-h session, two estimates of threshold were obtained for each of four to five conditions. Each listener was tested twice in counterbalanced order at each condition. The data presented in the figures represent the average of four estimates of threshold.

B. Results and discussion

Figure 1 contains the data obtained with the broadband masker. Thresholds (S/N in dB) are plotted as a function of the duration of the signal. The data points represent the average across three listeners and the error bars represent ± 1 standard error of the mean. Open symbols represent thresholds obtained with the NoSo configuration; closed symbols represent thresholds obtained with the NoSπ configuration. Squares indicate data obtained with no rove, triangles indicate data obtained while roving ITD, and diamonds indicate data obtained while roving IID.

We begin by discussing the data obtained in the standard, ''no rove'' conditions. As expected from previous findings, thresholds obtained in both the NoSo and NoSπ conditions increased as the duration of the signal decreased (e.g., Blodgett *et al.*, 1958; Green, 1966; Robinson and Trahiotis, 1972). They did not increase, however, in a parallel fashion. There was a relatively steep increase in NoSπ thresholds, as compared to their NoSo counterparts, as masker/signal duration was decreased from 135/125 to 35/25 ms. The data in Fig. 1 indicate that the masking-level differences (MLDs) decreased with duration and were 13.0, 11.5, and 5.5 dB for

masker/signal durations of 310/300, 145/125, and 35/25 ms, respectively. This outcome is consistent with data obtained earlier by Robinson and Trahiotis (1972) and by Yost (1985). The relatively greater loss in sensitivity in NoSπ conditions is known to be due to gating the masker on and off with the signal and is often referred to as the ''pulsed effect.'' The phenomenon was originally described by McFadden (1966). Robinson and Trahiotis (1972) and Yost (1985) showed that the relative loss of sensitivity in the NoSπ condition becomes more pronounced as duration is decreased.

We now turn to the thresholds obtained while roving either ITDs or IIDs. The data obtained in the NoSo condition (open symbols) indicate that roving the interaural parameters did not affect thresholds. To the degree that detection of the signal in the NoSo configuration is based on purely *monaural* cues, roving ITDs would not be expected to affect detection in the NoSo condition. This is so because an ITD only results in an absolute time delay of the noise or signal-plus-noise waveform in a monaural channel. We used independent samples of noise during each and every observation interval and estimated threshold by pooling information obtained across many such observation intervals. Consequently, absolute time delays, which only serve to determine the *particular* random sample of noise presented during a *particular* observation interval, could not affect our estimates of threshold. On the other hand, roving IIDs resulted in interval-by-interval variations in the overall level of the stimuli within a monaural channel. Such changes in the overall level of the stimuli have little, if any, affect on the ''monaural'' detectability of tonal signals in noise (e.g., Gilkey, 1987; Richards *et al.*, 1991; Kidd *et al.*, 1993). In summary, neither roving ITDs nor roving IIDs affected detectability in the NoSo condition, an outcome expected based on the assumption that detection in that condition depends on monaural rather than binaural information. At the same time the data indicate that the listeners were not ''distracted'' by interval-by-interval changes in the intracranial locus of binaural images produced by the stimuli.

The data obtained in the NoSπ condition (closed symbols) also indicate that roving the interaural parameters had little or no effect on thresholds. When the duration of the masker/signal was either 35/25 or 310/300 ms, the data obtained with and without roves are virtually superimposed on each other and the error bars for each condition overlap. When the duration of the masker/signal was 145/125 ms, the thresholds obtained with roving IIDs or roving ITDs were 1.4 and 2.7 dB greater, respectively, than the threshold measured in the absence of a rove.

The data in Fig. 1 were subjected to a three-factor, within-subjects analysis of variance. The error terms for the main effects and for the interactions were the interaction of the particular main effect (or the particular interaction) with the subject ''factor'' (Keppel, 1973). Consistent with visual evaluation of the data, the main effect of duration was a highly significant factor [$F(2,4)=488.19$, $p<0.001$, $\omega^2=0.55$] and accounted for 55% of the variability of the data. The main effect of signal type (So or Sπ) was also highly significant [$F(1,2)=932.11$, $p=0.001$, $\omega^2=0.35$] and accounted for 35% of the variability in the data. Although roving the stimuli did not produce a significant main effect [$F(2,4)=1.77$, $p=0.28$)], the interaction of rove by signal type was significant [$F(2,4)=8.35$, $p=0.037$, $\omega^2=0.004$] but only accounted for 0.4% of the variability in the data. This means that the differences in threshold produced by roving the NoSπ stimuli were relatively small, were not due to chance, but were larger than those obtained with the NoSo stimuli. At the same time however, those differences do not account for any substantial portion of the variability in the data. The duration by signal-type interaction was also significant [$F(2,4)=45.90$, $p=0.002$, $\omega^2=0.04$] and accounted for 4% of the variability in the data. This interaction reflects the smaller binaural release from masking obtained for short-duration stimuli when the masker is gated on and off with the signal. Neither the interaction of duration by rove, nor the triple interaction of duration by rove by signal type were statistically significant [$F(4,8)=1.44$, $p=0.31$ and $F(4,8)=0.66$, $p=0.636$, respectively]. In summary, the significant main effects and interactions accounted for about 94% of the variability in the data while indicating that the effects of roving ITDs or IIDs in the NoSπ condition were relatively small and only occurred in the NoSπ condition.

In order to evaluate the effects of roving ITD or IID on binaural processing *per se*, a second, two-way analysis of variance was performed only including the data obtained in the NoSπ condition. This analysis revealed significant main effects of duration [$F(2,4)=402.13$, $p<0.001$, $\omega^2=0.94$] and of rove [$F(2,4)=10.45$, $p<0.03$, $\omega^2=0.01$]. The duration by rove interaction was not significant [$F(4,8)=1.65$, $p=0.25$].

Because both analyses of variance indicate that roving the stimuli in the NoSπ condition affects performance, it is important to characterize and evaluate the effects of roving the stimuli in terms of their psychophysical salience. The largest effects of roving ITDs and IIDs occurred when the duration of the masker/signal was 145/125 ms. Roving IID increased the mean threshold by 1.4 dB, while roving the ITD increased the mean threshold by 2.7 dB in comparison to the mean threshold obtained in the no-rove condition. These increases in threshold can be cast in terms of increases in the standard deviations of the distributions of the interaural cues (Zurek, 1991; Bernstein and Trahiotis, 1995) that are produced by adding more intense Sπ signals to the diotic maskers in the conditions employing a rove. To perform the analysis, we assumed an effective binaural ''bandwidth'' of 250 Hz (a value in the upper range of various estimates of the binaural critical bandwidth centered at 500 Hz (e.g., Bourbon, 1966; Sever and Small, 1979; Hall *et al.*, 1983; Durlach *et al.*, 1986; Kohlrausch, 1988). In the no-rove condition, the average NoSπ threshold of -29 dB (displayed in Fig. 1), resulted in a distribution of ITDs having a standard deviation of 134 μs and a distribution of IIDs having a standard deviation of 2.8 dB. When IID was roved over the range of ± 6 dB, the NoSπ detection threshold increased by 1.4 dB. As a consequence, the underlying standard deviation of ITD increased by 19 μs and the underlying standard deviation of IIDs increased by 0.4 dB. When ITD was roved over the range of ± 500 μs, the threshold increased by 2.7 dB. This

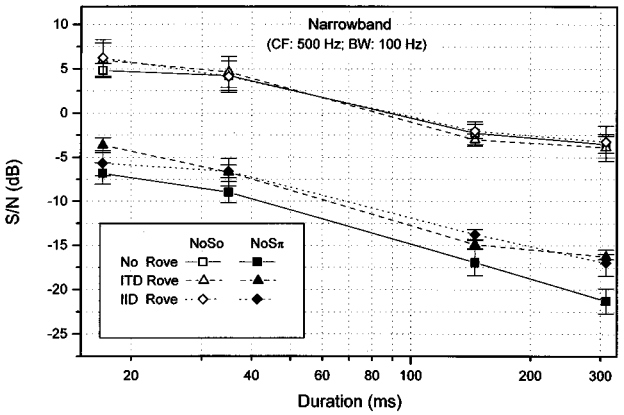


FIG. 2. Same as Fig. 1, but for the 100-Hz-wide masker centered at 500 Hz.

resulted in increases of the underlying standard deviations of ITDs and IIDs of 40 μs and 0.9 dB, respectively. These changes represent, in an absolute sense, extremely small effects. When considered relative to the magnitude of the ranges over which ITDs and IIDs were roved, the small changes in threshold that did occur suggest to us that the listeners effectively ignored or effectively overcame the interval-by-interval variability in the baseline values of the binaural cues. Stated in perceptual terms, the data suggest that the listeners were relatively unaffected by the interval-by-interval variability in the intracranial locus of the images produced by the noise-alone and noise-plus-signal waveforms, respectively. This means that informative changes in laterality are *not necessary* in order for threshold signal-to-noise ratios typical of those measured in a ''standard'' masking paradigm to be obtained.

Figure 2 contains the data obtained when the bandwidth of the masker was 100 Hz. The format of the figure is identical to that used in Fig. 1. Note that, as mentioned in Sec. I A, a masker/signal duration of 17 ms/12 ms was added. The trends in the data are qualitatively and statistically very similar to those obtained with the broadband masker (Fig. 1). Beginning with the no-rove condition, note that, once again, thresholds obtained in both the NoSo and NoSπ conditions increased as the duration of the signal decreased, albeit in a nonparallel fashion. Indeed, masking-level differences (MLDs) decreased with duration and were 17.7, 14.7, 13.2, and 11.7 dB for masker/signal durations of 310/300, 145/125, 35/25, and 17/12 ms, respectively. These decreases in the size of the MLD are smaller than those measured when the masking noise was broadband. The finding of a smaller ''pulsed effect'' with the narrow-band masker is consistent with data obtained earlier by Trahiotis (1973).

As observed with the broadband masker, the data obtained in the NoSo condition (open symbols) indicate that roving the interaural parameters did not affect thresholds. The data obtained in the NoSπ condition (closed symbols) indicate that roving the interaural parameters consistently produced increases in thresholds.

The data in Fig. 2 were subjected to the same three-factor, within-subjects analysis of variance discussed above. Consistent with visual evaluation of the data, the main effect of duration was a highly significant factor [$F(3,6)=74.11$, $p<0.001$, $\omega^2=0.32$] and accounted for 32% of the variability in the data. The main effect of signal type (So or Sπ) was also significant [$F(1,2)=76.66$, $p=0.013$, $\omega^2=0.58$] and accounted for 58% of the variability in the data. As before, although roving the stimuli did not produce a significant main effect [$F(2,4)=4.50$, $p=0.095$)], the interaction of rove by signal-type was significant [$F(2,4)=7.93$, $p=0.041$, $\omega^2=0.005$] but only accounted for 0.5% of the variability in the data. This interaction occurred because the differences in threshold produced by roving the NoSπ stimuli were larger than those obtained with the NoSo stimuli. The duration by signal-type interaction was *not* significant [$F(3,6)=2.90$, $p=0.124$] providing statistical confirmation of the weaker ''pulsed effect'' with the narrow-band masker. Neither the interaction of duration by rove, nor the triple interaction of duration by rove by signal type were statistically significant [$F(6,12)=1.24$, $p=0.354$ and $F(6,12)=1.11$, $p=0.41$, respectively]. In summary, the significant main effects and interactions accounted for about 95% of the variability in the data. The effects of roving ITDs or IIDs in the NoSπ condition were somewhat larger than those measured with the broadband masker.

As before, in order to evaluate the effects of roving ITD or IID on binaural processing *per se*, a second, two-way analysis of variance was performed only including the data obtained in the NoSπ condition. This analysis revealed significant main effects of duration [$F(3,6)=41.84$, $p<0.001$, $\omega^2=0.82$] and of rove [$F(2,4)=7.2$, $p=0.05$, $\omega^2=0.05$]. The duration by rove interaction was not significant [$F(6,12)=1.27$, $p=0.34$].

The effects of roving ITDs and IIDs in the NoSπ condition can, once again, be cast in terms of relative increases in the standard deviations of the distributions of the interaural cues that result from adding relatively more intense Sπ signals to the diotic maskers in the conditions employing a rove. To perform the analysis, we used a bandwidth of 100 Hz (the bandwidth of our narrow-band masker). This assumes that the effective binaural ''critical bandwidth'' is at least 100 Hz wide. Table I contains the standard deviations of the distributions of ITDs and IIDs, respectively produced by adding threshold-level Sπ signals to diotic maskers for the 12 experimental conditions. Again, the standard deviations were calculated from Zurek's (1991) equations and approximations. When the duration of the masker/signal was 310/300 ms, roving the ITD or IID resulted in 50%–60% increases in the standard deviation of the distributions of ITDs and IIDs. When the duration of the masker/signal was 145/125 ms, roving the ITD or IID resulted in 20%–35% increases in the standard deviation of the distributions of ITDs and IIDs. When the duration of the masker/signal was 35/25 ms, roving the ITD or IID resulted in about a 25% increase in the standard deviations of the distributions of ITDs and IIDs. Finally, when the duration of the masker/signal was 17/12 ms, roving the ITD or IID resulted in 10%–35% increases in the standard deviations of the distributions of ITDs and IIDs. Overall, the effects of roving either ITDs or IIDs were essentially equivalent. As found with the broadband masker, the increases in the widths of the distributions of ITDs and IIDs necessitated by roving the stimuli were

TABLE I. Standard deviations of the ITD and IID distributions, respectively, resulting from the addition of the $S\pi$ signal to the narrow-band diotic masker at threshold for each of the 12 experimental conditions (four durations×three rove types). Values were calculated using Zurek's (1991) equations and approximations.

	No rove		ITD rove		IID rove	
Duration (ms)	ITD (μs)	IID (dB)	ITD (μs)	IID (dB)	ITD (μs)	IID (dB)
17	409	8.6	557	11.8	457	9.7
35	332	7.0	413	8.8	417	8.8
145	154	3.3	188	4.0	209	4.4
310	101	2.1	164	3.5	153	3.3

very small in comparison to the range over which the respective interaural parameters were roved. Thus the data collected with the narrow-band masker suggest that the listeners were relatively unaffected by the interval-by-interval variability in the intracranial locus of the images of the stimuli. Overall, both sets of data suggest that changes in laterality, *per se*, are not *necessary* in order for listeners to exhibit ''normal'' binaural detection thresholds.

II. EXPERIMENT 2

The data from experiment 1 show that interval-by-interval changes in the intracranial loci of the stimuli produce only small degradations in listeners' abilities to detect an $S\pi$ signal added to diotic noise. This outcome suggests that listeners can detect the addition of the signal on the basis of changes in the widths of the distributions of the interaural cues, independent of their mean values. In order to test the generality of the findings from experiment 1, the same three listeners were tested in interaural correlation-discrimination conditions paralleling those employed for $NoS\pi$ detection in experiment 1.

A. Procedure

Listeners' sensitivities to interaural decorrelation of 100-Hz-wide narrow bands of noise centered at 500 Hz were measured utilizing a two-cue, two-alternative temporal forced-choice paradigm. In the absence of a rove, the first and fourth intervals of each trial contained diotic noise and served as cues. The listener's task was to detect an interaural decorrelation of the noise that occurred, with equal *a priori* probability, in either the second or third interval. The nonsignal interval also contained diotic noise. Three or four values of the interaural correlation were used to obtain a psychometric function. The value of the interaural correlation during the signal interval was fixed within a block of 80 trials. The values of correlation employed to generate each psychometric function were determined separately for each listener and each condition. The values of correlation were assigned in random order and testing was conducted twice at each value in a counterbalanced fashion. In this manner, two 80-trial blocks were run for each value of correlation tested. Sensitivity (d') at each of the values of correlation was evaluated by averaging d''s obtained in the two 80-trial blocks. For each condition, the points were fit by the equation:

$$\log d' = \log m + k \log(1-\rho), \tag{1}$$

where ρ is the interaural correlation of the narrow band of noise. By interpolation, the values of d' and ρ corresponding to 71% correct in a two-alternative task, i.e., the level of performance that defined threshold in experiment 1, were obtained.

In order to present a particular interaural correlation of the noise, we employed the well known and often used ''two-generator'' method of suitably scaling and mixing a pair of independent, Gaussian narrow-band noises. (e.g., Licklider and Dzendolet, 1948; Grantham and Wightman, 1979). Specifically, for each block of 80 trials, a pair of 1.6-s-long samples of noise having a desired interaural correlation was generated. Each pair of 1.6-s-long samples of noise was used to provide more than 50 temporally overlapping, but statistically independent tokens. The tokens of noise were selected randomly with replacement both within and across trials. The durations and rise/decay ramps of the bands of noise employed in this experiment were the same as those of the maskers employed in experiment 1. Roving ITDs and IIDs were employed as described for experiment 1. Roving the ITD alters the interaural correlation of the stimulus. However, because external delays can be compensated by internal delays, the *maximum* of the cross-correlation function, as well as its *distribution about the maximum* is unaltered by the roving ITD. Consequently, we will treat the baseline or ''reference'' correlations in the roving ITD condition as though they have a value of unity. In addition, it is assumed that roving IID has no effect on the computed value of correlation because, in accord with correlation-based models of binaural processing, all interaural correlations are ''normalized.'' For these reasons, the dependent variable is $\Delta\rho$, where ρ is determined by the ''two-generator'' mixing technique described above.

B. Results and discussion

Figure 3 displays threshold changes in interaural correlation ($\Delta\rho$) as a function of the duration of the stimuli (closed symbols). Error bars have been omitted for clarity. For 9 of the 12 conditions, the standard error was less than 0.015. The largest standard error was 0.0376, and was obtained at the duration of 17 ms for the condition in which ITD was roved. For comparison, the figure also contains the $NoS\pi$ thresholds from Fig. 2 transformed to $\Delta\rho$ using the relation: $\rho=(1-S/N)/(1+S/N)$ (e.g., Durlach *et al.*, 1986). As in Figs. 1 and 2, the data points represent the average across three listeners. Note that both the ordinate and the abscissa are scaled logarithmically. The direct measures of correlation discrimi-

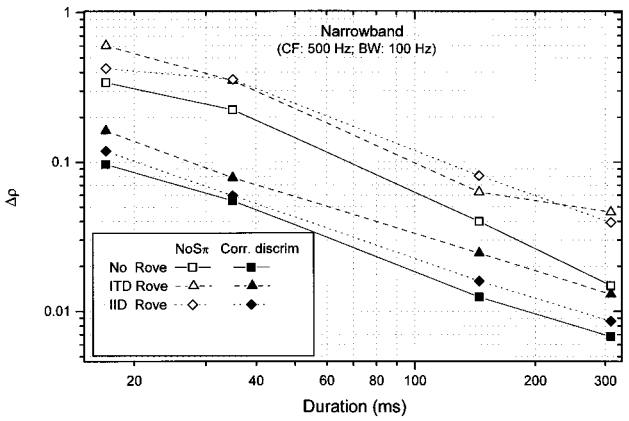


FIG. 3. $\Delta\rho$ as a function of duration. Closed symbols represent data obtained in the direct correlation discrimination task. Open symbols represent the NoSπ thresholds from Fig. 2 transformed to $\Delta\rho$. The data points represent the average across three listeners.

nation (closed symbols; solid lines) indicate that $\Delta\rho$ decreases proportionately with duration such that a doubling of duration leads to a halving of $\Delta\rho$. Consistent with the findings from experiment 1 (NoSπ detection), roving ITDs or IIDs led to small, but consistent reductions in sensitivity. A two-way analysis of variance using $\log_{10}(\Delta\rho)$ as the dependent measure revealed significant main effects of duration [$F(3,6)=36.50$, $p<0.001$, $\omega^2=0.81$] and of rove [$F(2,4)=13.91$, $p=0.016$, $\omega^2=0.03$]. The duration by rove interaction was not significant [$F(6,12)=0.20$, $p=0.97$].

A comparison of these correlation-discrimination data with the transformed NoSπ detection thresholds obtained in experiment 1 (open symbols) reveals that (1) sensitivity to changes in correlation was more acute in the direct correlation discrimination task and (2) the effects of roving ITDs and IIDs in the two conditions are very similar, with the possible exception of the data obtained when the duration was 310 ms.

Our finding that correlation thresholds obtained in the direct discrimination task were lower than those obtained in the NoSπ detection task was surprising because thresholds measured in the two types of task were found to correspond much more closely in two earlier studies in which similar comparisons were made (Koehnke *et al.*, 1986; Jain *et al.*, 1991). We considered whether our results could be attributed to our using the two-cue, two-alternative fixed procedure to measure direct discrimination thresholds while using a standard, two-alternative adaptive task to measure NoSπ detection thresholds. In order to evaluate this possibility, we examined other data obtained in our laboratory in an experiment with different listeners. In that experiment, we also obtained lower thresholds in the direct correlation-discrimination task even though fixed levels of the independent variables were used with the two-cue, two-interval paradigm to measure correlation-discrimination and NoSπ thresholds. In addition, with the listeners in the present study, we measured thresholds in both types of tasks using a single-interval procedure. Once more, lower thresholds were obtained with the direct correlation-discrimination task. Thus multiple sets of data indicate that the lower thresholds obtained with the direct discrimination task shown in Fig. 3 were not due to the use of different psychophysical procedures to measure the two types of thresholds. Rather, the differences in threshold appear to reflect true differences in sensitivity.

We then considered whether our listeners had relatively high NoSπ detection thresholds and/or relatively low correlation-discrimination thresholds as compared to the listeners who participated in the studies of Koehnke *et al.* (1986) and Jain *et al.* (1991). Scrutiny of the data from each of the studies indicated that this was not so. There was a great deal of overlap in the thresholds obtained across the studies both within and across the two types of measures. Consequently, at this time we can offer no explanation for the differences in the findings across studies.

Using computer simulations, we found that sample-to-sample variations in the correlation were *greater* when a ''target'' value of correlation was produced by mixing a pair of independent sources of noise as compared to when that same ''target'' value of correlation was produced by adding an antiphasic tone to a diotic band of noise (NoSπ). In future investigations, we plan to determine whether differences in the sampling distributions of interaural correlation can account for our listeners' having different thresholds, each defined in terms of $\Delta\rho$, in the two types of psychophysical tasks. Accounting for the differences in sensitivity with the two tasks is beyond the scope of this investigation, which concerns the effect of roving ITDs or IIDs on NoSπ and direct correlation-discrimination thresholds.

III. GENERAL DISCUSSION

The results of experiments 1 and 2 indicate that roving ITDs or roving IIDs on an interval-by-interval basis leads to only small increases in threshold obtained in both NoSπ detection and interaural correlation-discrimination tasks. The small magnitudes of the increases in threshold due to roving ITDs or IIDs, when considered in the context of the large range over which the interaural cues were roved, support the conclusion that *informative changes in laterality* of the intracranial images of the stimuli are *not required* in order for detection to occur. Instead, it appears that listeners are able to detect on the basis of changes in the widths of the distributions of ITDs and IIDs produced by adding the signal to the masker. Such changes are independent of changes in laterality produced by roving ITDs or IIDs across observation intervals.

It should be noted that the changes in the widths of the distributions of ITDs and IIDs produced by adding the signal to the masker are probably manifested perceptually in a manner dependent on the bandwidth of the masker. For narrowband maskers, listeners often report that the perceptual cue they use is a ''broadening'' of the intracranial image when the signal is added to the masker, as compared to the more punctate intracranial image produced by the masker presented in isolation. For broadband maskers, listeners virtually always report that the perceptual cue they use to detect the signal is the occurrence of a ''tonal'' or ''pitchlike'' quality that does not exist when the masker is presented in isolation.

Overall, our data appear to be consistent with those reported by Richards and Henning (1994) at the 127th meeting of the Acoustical Society of America. They reported that roving IIDs across observation intervals in a standard NoSo/NoSπ detection task generally reduced MLDs by ''small amounts.'' To our knowledge, these are the only data available for purposes of comparison.

As mentioned above, we also studied the effects of roving ITDs and IIDs in a supplementary experiment using a single observation interval task. Our purpose in using the single interval task was twofold: (1) it allowed us to remove within-trial information regarding the reference stimuli so that ''memory concerning the reference stimuli'' had to be accrued *across* trials, and (2) it allowed us to partition the data conditional on particular values of ITD or IID to determine whether deleterious effects due to roving the interaural disparities were, in turn, due to large baseline ITDs or IIDs.

The effects due to roving ITDs and IIDs within the single interval task were essentially equivalent to those obtained in the main experiments using tasks with multiple intervals. Furthermore, thresholds did not depend on particular values of ITD or IID. This indicates that the small deleterious effects due to roving the interaural disparities probably result from forcing the listeners to detect or to discriminate within the *context* of roving ITDs or IIDs.

It occurred to us that it was logically possible for the listeners to make themselves effectively immune to the rove if they were somehow able to access, *independently*, internal information regarding ITDs and IIDs. If that were true, the listener could focus on changes in ITD (or, symmetrically, IID) when the physical dimension of the rove was IID (or, symmetrically, ITD), a strategy that would drive performance toward the ''no-rove'' condition. We reasoned that such a strategy could be defeated by roving *both* ITDs and IIDs simultaneously. In order to test this, we reran all three listeners in both tasks using the 310-ms duration. Roving both ITDs and IIDs simultaneously produced deficits not unlike those produced by roving either disparity alone. Therefore, it appears reasonable to conclude that the listeners did not avoid the effects of the rove in the main experiments by having independent access to ITDs or IIDs.

The ability to detect and discriminate changes in the distribution of binaural disparities independent of their ''reference'' or baseline value is, in fact, consistent with modern models of binaural processing. For example, the data can be accounted for via Durlach's equalization-cancellation model (e.g., Durlach, 1963, 1972) by incorporating his proposal that multiple equalization transformations could be accomplished in parallel. Recently, Schneider and Zurek (1989) successfully used Durlach's EC model to account for their lateralization data by incorporating the assumption that cancellation could be accomplished, in parallel, at each of a number of simultaneously available values of internal delay. In order to account for the small effects of roving ITDs *and* IIDs in our experiments, one need only assume that the listeners could also equalize any of several differences of interaural intensity in a parallel fashion. The assumption is that decisions would be made using the compensatory ITD or IID that minimizes the external noise masker (i.e., maximizes signal-to-masker ratio).

Our data can also be accounted for by cross-correlation based models (e.g., Colburn, 1973; Stern and Colburn, 1978) which utilize changes of the patterns of activity to account for detection and discrimination data. In this context, roving ITDs or IIDs would merely alter the left/right displacement of the activity from midline without affecting the local properties of the distribution (i.e., shape) of the activity. This approach provides a mechanistic framework which permits the perceptual laterality of the stimulus to be varied by roving ITD or IID while simultaneously permitting detection or discrimination to occur on the basis of local properties such as the width or diffuseness/compactness of the intracranial image.

IV. SUMMARY

The purpose of this study was to assess whether roving baseline interaural parameters of the stimuli would adversely affect performance in masking-level difference and in interaural correlation-discrimination paradigms. Overall, the results indicate that roving the interaural cues produced little, if any, degradations of performance as compared to data obtained in the absence of a rove. This outcome occurred even with stimuli of the shortest duration. In terms of the perceptual aspects of the task, the data indicate that interval-by-interval variation of the lateral position of the stimuli (which were heard in one of many positions from far toward the left ear to far toward the right ear) did not affect performance. Although not intuitively obvious, this ''negative'' outcome is exactly what one would expect from modern models of binaural processing.

ACKNOWLEDGMENTS

The authors wish to thank Dr. Joe Hall III, Dr. Wes Grantham, and one anonymous reviewer for their helpful comments. Supported by Research Grant Nos. NIH DC-02103 and NIH DC-00234 from the National Institute on Deafness and Other Communication Disorders, National Institute of Health.

Bernstein, L. R., and Trahiotis, C. (**1995**). ''Binaural interference effects measured with masking-level difference and with ITD- and IID-discrimination paradigms,'' J. Acoust. Soc. Am. **98,** 155–163.

Blodgett, H. C., Jeffress, L. A., and Taylor, R. W. (**1958**). ''Relation of masked threshold to signal-duration for various interaural phase-combinations,'' Am. J. Psychol. **71,** 283–290.

Bourbon, W. T., Jr. (**1966**). ''Effects of bandwidth and level of masking noise on detection of homophasic and antiphasic tonal signals,'' dissertation, University of Texas, Austin, TX.

Colburn, H. S. (**1973**). ''Theory of binaural interaction based on auditory-nerve data I. General strategy and preliminary results on interaural discrimination,'' J. Acoust. Soc. Am. **54**, 1458–1470.

Durlach, N. I. (**1963**). ''Equalization and cancellation theory of binaural masking-level differences,'' J. Acoust. Soc. Am. **35,** 1206–1218.

Durlach N. I. (**1972**). ''Binaural signal detection: Equalization and cancellation theory,'' in *Foundations of Modern Auditory Theory*, edited by J. V. Tobias (Academic, New York), Vol. II.

Durlach, N. I., Gabriel, K. J., Colburn, H. S., and Trahiotis, C. (**1986**). ''Interaural correlation discrimination: II. Relation to binaural unmasking,'' J. Acoust. Soc. Am. **79,** 1548–1557.

Gilkey, R. H. (**1987**). ''Spectral and temporal comparisons in auditory masking,'' in *Auditory Processing of Complex Sounds*, edited by W. A. Yost and C. S. Watson (Erlbaum, Hillsdale, NJ).

Grantham, D. W., and Wightman, F. L. (**1979**). ''Detectability of a pulsed tone in the presence of a masker with time-varying interaural correlation,'' J. Acoust. Soc. Am. **65,** 1509–1517.

Green, D. M. (**1966**). ''Interaural phase effects in the masking of signals of different durations,'' J. Acoust. Soc. Am. **39,** 720–724.

Hafter, E. R., Bourbon, W. T., Blocker, A. S., and Tucker, A. (**1969**). ''A direct comparison between lateralization and detection under conditions of antiphasic masking,'' J. Acoust. Soc. Am. **46,** 1452–1457.

Hall, J. W., Tyler, R. S., and Fernandes, M. A. (**1983**). ''Monaural and binaural auditory frequency resolution measured using bandlimited noise and notched-noise masking,'' J. Acoust. Soc. Am. **73,** 894–898.

Henning, G. B. (**1973**). ''Effect of interaural phase on frequency and amplitude discrimination,'' J. Acoust. Soc. Am. **54,** 1160–1178.

Jain, M., Gallagher, D. T., Koehnke, J., and Colburn, H. S. (**1991**). ''Fringed correlation discrimination and binaural detection,'' J. Acoust. Soc. Am. **90,** 1918–1926.

Keppel, G. (**1973**). *Design and Analysis: A Researchers Handbook* (Prentice-Hall, Englewood Cliffs, NJ).

Kidd, Jr., G., Mason, C. R., Brantley, M. A., and Owen, G. A. (**1989**). ''Roving-level tone in noise detection,'' J. Acoust. Soc. Am. **86,** 1310–1317.

Koehnke, J., Colburn, H. S., and Durlach, N. I. (**1986**). ''Performance in several binaural-interaction experiments,'' J. Acoust. Soc. Am. **79,** 1558–1562.

Kohlrausch, A. (**1988**). ''Auditory filter shape derived from binaural masking experiments,'' J. Acoust. Soc. Am. **84,** 573–583.

Levitt, H. (**1991**). ''Transformed up-down methods in psychoacoustics,'' J. Acoust. Soc. Am. **49**, 467–477.

Licklider, J. C. R., and Dzendolet, E. (**1948**). ''Oscillographic scatterplots illustrating various degrees of interaural correlation,'' Science **107,** 121–124.

McFadden, D. M. (**1966**). ''Masking-level differences with continuous and with burst masking noise,'' J. Acoust. Soc. Am. **40,** 1414–1419.

Richards, V. M., Heller, L. M., and Green, D. M. (**1991**. ''Detection of a tone in a narrow band of noise: The energy model revisited,'' Q. J. Psychol. **43A**, 481–503.

Robinson, D. E., and Trahiotis, C. (**1972**). ''Effects of signal duration and masker duration on detectability under diotic and dichotic listening conditions,'' Percept. Psychophys. **12,** 333–334.

Schneider, B. A., and Zurek, P. M. (**1989**). ''Lateralization of coherent and incoherent targets added to a diotic background,'' J. Acoust. Soc. Am. **86,** 1756–1763.

Sever, J. C., and Small, A. M. (**1979**). ''Binaural critical masking bands,'' J. Acoust. Soc. Am. **66,** 1343–1350.

Stern, R. M., and Colburn, H. S. (**1978**).'' Theory of binaural interaction based on auditory-nerve data. IV. A model for subjective lateral position,'' J. Acoust. Soc. Am. **64**, 127–140.

Trahiotis, C. (**1973**). Unpublished report.

Yost, W. A. (**1985**). ''Prior stimulation and the masking-level difference,'' J. Acoust. Soc. Am. **78,** 901–907.

Zurek, P. M. (**1991**). ''Probability distributions of interaural phase and level differences in binaural detection stimuli,'' J. Acoust. Soc. Am. **90,** 1927–1932.

Characteristics of a pulsating jet through a small modulated orifice, with application to voice production

Luc Mongeau and Nancy Franchek
School of Mechanical Engineering, Purdue University, 1077 Ray Herrick Laboratories, West Lafayette, Indiana 47907-1077

Cecil H. Coker and Robert A. Kubli
Multimedia Communications Laboratory, Bell Laboratories, Lucent Technologies, 700 Mountain Avenue, Murray Hill, New Jersey 07974

(Received 25 September 1996; accepted for publication 15 April 1997)

A detailed understanding of the aerodynamics of air flow in the larynx and the vocal tract is needed to refine physiological models of human voice production. This understanding can be applied in speech synthesis, voice diagnostics, and voice recognition. To date, most aeroacoustic models of phonation have been based on Bernoulli's orifice theory, i.e., the assumption that flow phenomena within the larynx are ''quasi-steady.'' This assumption, however, has never been rigorously verified experimentally. In this study, detailed aerodynamic measurements were performed of a pulsating open jet through a modulated orifice with a time-varying area. Orifice geometry and characteristic Reynolds numbers and Strouhal numbers of the pulsating jet flow were representative of speech production. Simple source-filter models based on the quasi-steady flow assumption and an ideal one-dimensional monopole source model were found to yield satisfactory velocity, flow rate, and dynamic pressure predictions for most of the duty cycle. Significant deviations from quasi-steady behavior were observed only during the early part of the duty cycle, where the flow velocity in the center core rapidly reached a peak value immediately after release of the false folds. This acoustic near-field phenomenon did not affect the pressure waves generated by the pulsating jet through the orifice, propagated in the long rigid tube upstream of the orifice. The impact on this phenomenon on actual sound generation within the larynx, and wall pressures on the vocal folds, is not clear. [S0001-4966(97)02408-9]

PACS numbers: 43.70.Aj, 43.28.Ra [AL]

INTRODUCTION

Phonation is the mechanism of human voice production. It is used in speech, for vowels and voiced consonants (Flanagan, 1965). It is the essential part of singing. Progress in many areas of speech sciences relies on our capacity to understand, model, and reproduce the process of phonation. Problems in vocal fold morphology such as early nodules, cysts, granulomas, or cancer may be nonintrusively detected and diagnosed by recording the changes occurring in a patient's voice (Scherer, 1991a). Possible consequences of phonosurgery on a patient's ability to communicate may be evaluated (Scherer, 1991b). Speech synthesizers based on physiological models of speech production may be used to assist patients suffering from voice disorders. Other applications requiring a fundamental understanding of pulsating flows through modulated orifices include intra-cardiac jets (Diebold *et al.*, 1990), and also sprays, pulsed combustion, pulse-tube refrigeration, sirens, and various musical instruments (for example, Fletcher, 1979).

The basic mechanics of phonation are well understood, and have been explained by Wegel (1930), Flanagan (1965), Titze (1973, 1974), Cranen (1987), and others. The source of phonation is located within the larynx. Phonation is initiated when muscles compress the rib cage, thereby increasing lung pressure and expelling air through the trachea into the pharynx. The vocal folds, two lips of ligament and muscle at the top of the trachea, are then brought together by muscular action to form a constriction in the wind pipe. When a certain critical lung pressure is achieved, self-sustained oscillations of the vocal folds are initiated. A pulsating jet is formed by the air flowing through the glottis, a time-varying orifice between the vocal folds. Periodically, lung pressure buildup forces the glottis to open. The decrease in pressure accompanying the formation of the jet together with the vocal folds compliance then force the glottis to close, and the cycle repeats. Typical values for lung pressure, air flow rate, and other flow properties can be found, for example, in Flanagan (1965).

The mechanism for self-oscillation of the vocal folds has many characteristics of the oscillations of other similar physical systems such as collapsible tubes (Berke *et al.*, 1991), reed musical instruments (Hirschberg *et al.*, 1990), and tongue-tip trills (McGowan, 1992). A pulsating flow is established through a modulated orifice with deformable, compliant walls. Approximate analytical models for these systems are based on electroacoustical analogies (Kathuriya and Munjal, 1979; Beranek, 1986; Allen and Strong, 1985; Koizumi *et al.*, 1985; Anathapadmanabha and Fant, 1982; and others). The Norton equivalent electrical circuit can be determined in which the periodic flow through the modulated orifice is represented by a current source with an internal shunt impedance, and the surrounding pipes and cavities are represented as an all-pole filter (Flanagan, 1965). Since the

internal impedance of the source varies with time (as the geometry varies during one oscillation period), the source and the filter are not linearly separable, but the process constitutes a nonlinear feedback loop. Computer predictions of the output can be made using time-marching algorithms (see, for example, Gupta and Schroeder, 1993).

Although a small fraction of the sound emitted by the vibrations of the vocal cords may be radiated through structural vibrations of the bones and cartilage, most of the acoustic energy is generated from the gas pulsations in the vocal tract. These flow pulsations are converted into acoustic waves that propagate into the vocal tract and are radiated from the lips and the nasal cavity. The mechanism for converting the pulsating jet into sound, i.e., the periodic injection of air mass into the vocal tract, is ordinarily modeled as a classic one-dimensional monopole sound source in a duct (Pierce, 1989; Davies and Ffowcs Williams, 1968). The acoustic source strength is consequently assumed to be directly proportional to the unsteady mass flux through the orifice.

In speech, semi-empirical models for the behavior of the human glottis have been suggested by Fant (1982), Flanagan and Landgraf (1968), Ishizaka and Flanagan (1972), Titze and Talkin (1979), and others. Knowledge of the so-called non-linear ''glottal impedance,'' $Zg(t)$, is essential to model the glottal oscillation mechanism and the vocal tract acoustic properties (for example formant frequencies). The ''glottal impedance,'' unrelated to the familiar acoustical or mechanical impedance, characterizes the pressure-flow relationship within the glottis, i.e., it is defined as the ratio of the transglottal pressure head and the mass flow rate through the orifice (van den Berg *et al.*, 1957). It is a time-varying quantity with both steady (dc) and periodic (ac) components. The steady component is due to the fact that the real glottis does not completely shut; there remains a small opening after closure between the arytenoid cartilages.

Due to the complexity of pulsating jet flows, which involve turbulence and flow separation, and that of the glottal airway geometry, it is impossible to calculate the glottal impedance exactly using analytical or numerical methods. [Recent efforts in this direction were made by Alipour *et al.* (1996).] Direct measurements *in vivo* are also difficult to perform and do not allow the control of the glottis geometry, which makes it necessary to take the vocal folds motion into consideration. Despite progress in vocal tract imaging using, for example, magnetic resonance imaging (MRI), detailed measurements of the time-varying geometry of the glottal airway in human subjects are difficult. Furthermore, ''*in vivo*'' flow measurements are also difficult, and may be hazardous.

In a first category of glottal impedance investigations, studies have been made using static configurations that replicate the geometry of the human vocal apparatus. The instantaneous impedance was then assumed to be equal to the steady (dc) impedance of a geometrically similar system at rest, i.e., the approximation was made that flow phenomena were quasi-steady. Static pressure-flow relationships have been obtained experimentally by Wegel (1930) using crude mechanical models. Later, van den Berg *et al.* (1957) arrived at an expression for the glottal impedance based on experiments using an excised dog larynx. Further work on static configurations was performed using mechanical models by Ishizaka and Matsudaira (1972), Scherer (1981), Scherer *et al.* (1983), Binh and Gauffin (1983), and Scherer and Guo (1990). The steady glottal impedance and detailed information about steady orifice flows may also be obtained computationally, using for example the finite element method (Guo and Scherer, 1993). The advantage of steady models lies in their relative simplicity, due to the neglection of dynamic flow phenomena. This approach could lead to detailed flow simulations for complex orifice shapes that would require much less computational effort than truly dynamic simulations involving unsteady flow models. The underlying assumption, however, needs to be verified for a representative range of glottal operating conditions.

A second category of investigations addressed the problem of the acoustic impedance of fixed orifices. In these studies, the upstream pressure was modulated by driving the system with a sound source, such as a loudspeaker, thereby creating a pulsating jet. The acoustic impedance of fixed orifices has been measured by Ingàrd and Labate (1950), Ingàrd (1953), Laine and Karjalainen (1986), and Rösler and Strube (1989) using a geometry closer to that of the glottis. Detailed aerodynamic measurements of modulated jets through static orifices have also been performed by Iguchi *et al.* (1990), Ikeda *et al.* (1994), Pelorson *et al.* (1994, 1995, 1996), and others. Although pulsating jets through fixed orifices are very close to the actual flow field within the human glottis, such flows may not reproduce some important effects due to the periodic blockage of the orifice, the resulting subglottal pressure buildup, and the impulsively created jet at the moment of opening. Berke *et al.* (1989) first measured the particle velocity at the discharge of an excised canine larynx using a hot-wire anemometer. Using a similar approach, Alipour and Scherer (1995) observed significant flow nonuniformity in the pulsating jet velocity profile. However, these results may have been caused by the difficulty in excised larynx studies to maintain complete control over the orifice geometry and to position the hot-wire probe very close to the orifice discharge, in the potential core of the jet.

Shadle *et al.* (1987) investigated a mechanically modulated orifice with a time-varying area, but the scope was limited to flow visualization and static impedance measurements. There has been no detailed fluid dynamic studies of pulsating jets through mechanically modulated orifices in the range of pressures and flow rates encountered in speech, although closely related studies may be found in the literature. The initial stages of jet formation in a flue organ pipe, for example, was investigated by Verge *et al.* (1994). Kozenko *et al.* (1989) investigated the characteristics of a pulsating jet produced by the transverse vibrations of a coaxial cylinder. The properties of viscous jets with large velocity modulations were examined by Bousfield *et al.* (1990) for applications to ink jet printers. These investigations, however, provide little useful information for glottal flow and glottal impedance modeling.

Some attempts have also been made to measure the glottal impedance *in vivo*. Sondhi (1975) measured the wave-

form of the unfiltered voice by placing a long anechoic tube close to a human glottis. Cranen (1987) performed extensive measurements of the transglottal pressure and the glottal area, from which the mean lung pressure and the glottal flow rate was deduced. Neither study, however, provided sufficiently detailed information about the flow field in the glottal region. Many questions were raised by Cranen (1987) regarding the detailed features of laryngeal flow. Such inquiries can also be found in McGowan (1988), Kaiser (1983), and Teager and Teager (1983).

The question of the range of validity of the quasi-steady assumption is extremely important for the modeling of phonation, as emphasized by McGowan (1993). Building accurate static mechanical models is relatively easy compared to models with time-varying orifice geometries. There are many detailed aerodynamic studies of open jet flows (Schlichting, 1979), and confined steady jets (So *et al.*, 1987). Comparable studies of pulsating jets, however, are very sparse. Iguchi *et al.* (1990) recently compared the properties of steady open jets and pulsating open jets through orifices with constant area. They reported that the jet velocity profile was significantly affected by the acceleration and the deceleration of the flow field. The pulsating jet was more widely spread during the deceleration phase than during the acceleration phase. The growth rate of the boundary layer in the developing flow region was greater during the deceleration phase than during the accelerating phase. These results were obtained for confined jets with flow velocities up to 30 m/s, an orifice diameter of 51 mm, and tube diameters of 78 and 150 mm. Diebold *et al.* (1990) performed a detailed study of pulsating confined hydraulic jets, including precise measurements of velocity profiles and flow patterns. Again, deviation from quasi-steady behavior was observed. McGowan (1993) discussed the question and concluded that the quasi-steady assumption may not always be valid in speech, based on studies of nonsteady channel flows, such as those of Pedley (1983), Pedley and Stepanoff (1985), and others. Experiments that are specifically designed for speech applications, however, are still needed in order to determine the range of validity of the quasi-steady assumption. Although acoustical excitation by pulsating jets occurs in a large number of engineering applications, such as sirens or pulsed combustion, there have been few fundamental studies of the excitation of plane waves in tubes by small pulsating jets though time-varying orifices, and none that conclusively establishes the validity of the quasi-steady hypothesis for the modeling of speech production.

The main purpose of this study was therefore to verify the quasi-steady assumption for a pulsating open jet flowing through a modulated converging orifice with a time-varying area, within a range of Strouhal numbers and Reynolds numbers typical of speech production. Another goal was to verify the accuracy of the one-dimensional monopole model for predicting the sound pressure generated by the pulsating jet, propagated as plane waves in the subglottal tube upstream of the glottal orifice. Finally, detailed flow measurements of the flow through the orifice were made to provide some information about the fluid dynamic characteristics of the pulsating jet flow, emphasizing dynamic effects, i.e., features that do not obey steady flow relationships.

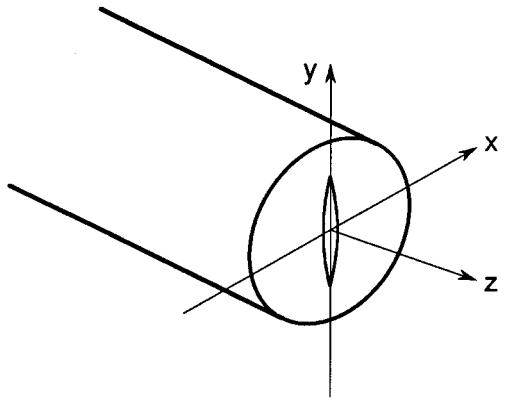


FIG. 1. Schematic of the rigid tube, the orifice, and the coordinate system.

I. THEORETICAL BACKGROUND

A. Steady-state relationships for flow through a fixed orifice

Consider a steady laminar flow through a straight rigid tube terminated by a small orifice, similar to that illustrated in Fig. 1. Assume the tube to be infinitely long in the negative z direction, with the z-coordinate axis coinciding with the center of the tube. Assume the flow downstream of the orifice, located in the center of the tube at $z=0$, to be a steady, open turbulent jet. Bernoulli's equation can be obtained by integrating the momentum equation along the streamline along the z axis. For a small orifice-to-tube-area ratio, the momentum of the fluid within the tube is negligible, leading to the well-known pressure-flow relationship

$$p_{0t}-p_{\text{atm}}=\tfrac{1}{2}\rho_0 U_{0c}^2, \tag{1}$$

where ρ_0 is the ambient density (kg/m^3) (assumed to be uniform), U_{0c} is the centerline flow velocity at the orifice discharge (m/s) (within the potential core of the jet), and $\Delta p_0=p_{0t}-p_{\text{atm}}$ is the static pressure gradient across the orifice (Pa). The flow velocity is not perfectly uniform across the orifice. Because of viscous effects, boundary layers are formed near the orifice walls. Flow entrainment and mixing causes the net mass flux of the jet to increase in the streamwise direction, and the jet to diffuse and the velocity profile to broaden. The volumetric flow rate, Q_0 (ml/s), must therefore be obtained by integrating the mean flow velocity immediately downstream of the orifice over the whole area of the orifice, A_g (m^2):

$$Q_0=\int\int_{A_g} U_0(x,y)\,dA_g\,. \tag{2}$$

Because of the effects of viscous stresses in the proximity of the walls, Eq. (1) is not satisfied exactly except towards the center of the jet. To account for this, Bernoulli's orifice theory expresses the orifice flow rate as a function of the pressure gradient:

$$Q_0=CA_gU_{0c}\,, \tag{3}$$

where the nondimensional orifice coefficient $C(\mathrm{Re},A_g/A_t)$ is given by the ratio of the space-averaged (indicated by an overbar) flow velocity and the maximum flow velocity at the center of the orifice:

$$C=\frac{\overline{U_0}^{xy}}{U_{0c}}. \tag{4}$$

In general, C is a function of the flow Reynolds number, the orifice geometry, and the ratio A_g/A_t of the orifice area and the tube area (White, 1986).

B. Unsteady flow relationships

Consider now the case where the orifice area $A_g(t)$ is a function of time. The quasi-steady assumption implies that, at every instant, the flow through the modulated orifice should be similar to that through a fixed orifice of the same geometry at that time. The pressure inside the tube may be decomposed into the sum of a mean value, p_{0t}, and a time-varying (acoustic) component, p', associated with acoustic waves (assumed to be plane waves) generated by the fluctuating mass flux at $z=0$ and propagated in the negative z direction, down the tube. It is convenient to define the mean pressure p_{0t} as the ''blocked'' pressure in the tube, i.e., the average pressure in the tube while the orifice is closed.

Since no wave reflections are assumed, the fluctuating pressure p' at any location down the tube z can be calculated assuming an ideal one-dimensional monopole source, using the relationship (Pierce, 1989)

$$p'(z,t)=-\frac{\rho_0 c}{A_t(t)}\,Q\left(t+\frac{z}{c}\right), \tag{5}$$

where c is the speed of sound (m/s) and Q is the flow rate evaluated at the retarded time $t+z/c$ (see, for example, Dowling and Ffowcs Williams, 1983).

Equation (5) can be expressed, accordingly with the quasi-steady assumption, in terms of the instantaneous values of $C(t)$, $A_g(t)$, and $\Delta p(t)=p_{0t}-p_0+p'$. A solution for Q and p' can then be obtained in terms of the mean pressure gradient and the (assumed known) orifice area by solving Eq. (5) together with Eq. (4) and one relationship $C=C(Q,A_g)$ which must be established empirically (see Sec. III A).

II. EXPERIMENTAL METHODOLOGY

The mechanical model of the human larynx was designed and built to reproduce the essential features of laryngeal aerodynamics, but with simplified acoustic boundary conditions, in order to allow a direct comparison between experimental results and the theoretical model described in the previous section. A sketch of the experimental apparatus is shown in Fig. 2. The orifice plate was a rubber model of a narrow section of the vocal tract that included the vocal folds and the glottal orifice. The nominal larynx model dimensions, shown in Fig. 3, were typical of those of a human larynx (Flanagan, 1965). The model was cast using a liquid rubber with a room temperature vulcanized (RTV) catalyst. The center of a 15-cm-long Plexiglas cylinder, having a 2.54-cm diameter, was first machined and polished to the dimensions of the glottal orifice in a rectangular Plexiglas

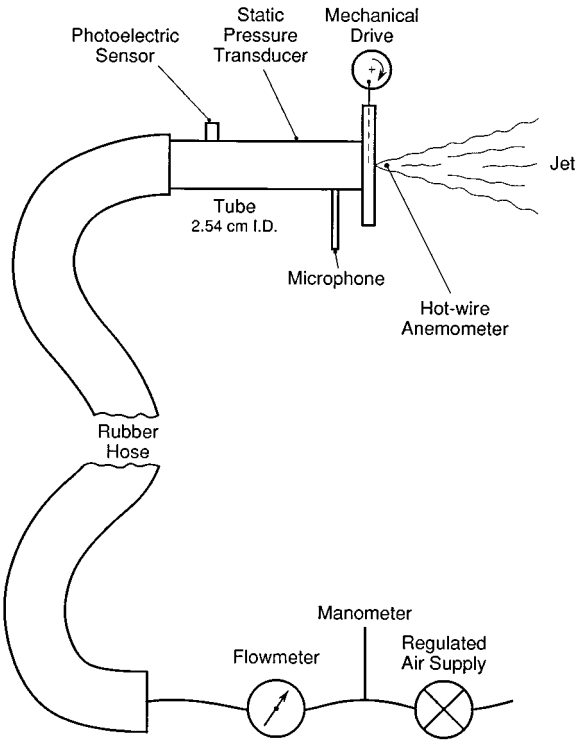


FIG. 2. Schematic of the experimental apparatus.

mold. The mold, having 10 cm×10 cm×1 cm inside dimensions, hosted the negative through centered 2.54-cm-diam holes through the top and the bottom. Before filling, two cross-shaped stainless steel actuating rods were inserted through each side of the mold. The resulting rubber model had the shape of the larynx fully opened. The laryngeal orifice could be closed by pressing the actuators, which deformed the rubber and brought the folds tightly together. Note that rubber deformation (from stress and strain within the material) may have induced some variability in the exact orifice geometry during one cycle. For example, the line formed by the point of closure may not have been exactly centered, or glottal opening may not have occurred everywhere simultaneously.

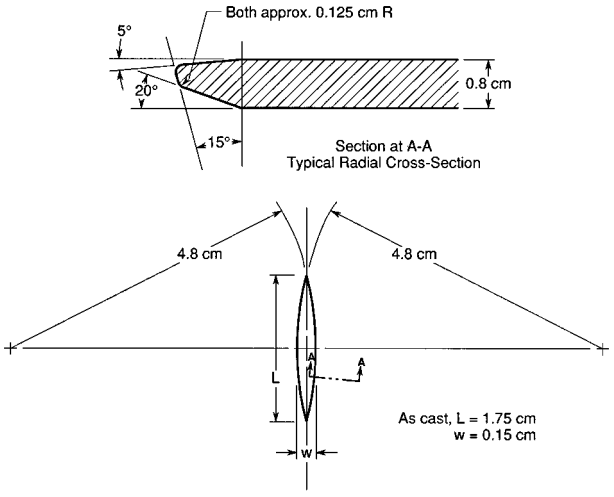


FIG. 3. Orifice geometry.

The larynx rubber model (or orifice plate) was clamped between two 6-mm-thick aluminum plates. Holes aligned with the laryngeal airway, 3.18 cm in diameter, allowed tubes to be installed upstream and downstream of the orifice. The vocal folds actuators, emerging from each side, were driven by eccentrics, electric dc motors, and belt-pulley assemblies. The rotational speed could be varied in the range $10 \leqslant \Omega \leqslant 150$ Hz. A custom-made feedback controller was used to synchronize the rotation of the two motors and regulate their speed. A tachometer mounted on one of the motor shafts allowed the rotational speed to be read on an LCD display and provided a periodic signal output to the data acquisition system. The assembly was mounted on a rigid aluminum frame, together with the motors and the driving system. The resulting apparatus allowed accurate dynamic control of the orifice geometry. Slippage of some rotating and translating parts, however, together with relaxation of the rubber material, caused the maximum area of the glottal opening to drift (between 14 and 18 mm^2, unpredictably). A monitoring procedure for the glottal area was therefore implemented. A photoelectric sensor located inside the tube provided an electric signal which was determined to be proportional to the orifice area (see Sec. III C below). The maximum area, $A_{g\,\max}$, was within about 3% of 15 mm^2 for most of the dynamic measurements reported in the following sections.

Upstream of the larynx, a 15-cm-long copper tube with an inside diameter of 2.54 cm was connected to a regulated compressed air supply through a 6-m-long, flexible rubber tube. Both extremities of the rubber tube were filled with mineral wool in order to minimize sound reflections from the tube end and to reduce extraneous noise from the air supply. The transition between the orifice plate and the subglottal tube was abrupt; no attempt was made to replicate the actual transition between the trachea and the vocal folds, which is more gradual. The mean flow rate through the air supply was measured using a precision mass-flow meter (Baratron type 558A) and a digital voltmeter. The time-averaged pressure drop across the laryngeal orifice, $\overline{\Delta p'}$, was measured using a precision manometer (Baratron type 698A). The low-pressure port of the transducer was open and the high-pressure port was connected to a pressure tap located 2.5 cm upstream of the glottal plane (the open jet discharged directly to atmosphere). Static pressure and air temperature were also measured at the flow supply, immediately downstream of the regulating valve. A dual channel anemometer system (TSI type IFA 100) with hot-wire probes (TSI model 1210-T1.5) and probe supports (TSI model 1150) were used for the flow velocity measurements. A 3.8-μm-diam platinium-coated tungsten wire was used, with a sensing element length of 1.27 mm, and a spacing between the wire supports of 1.5 mm. The wire was oriented along the y direction (vertically). The hot-wire probes were traversed in the xy plane using two orthogonal, manually actuated micrometric position control devices. The sensing element was located as close as physically possible to the orifice, avoiding contact between the rubber folds (which bulged out slightly when forced together by the actuators) and the very fragile hot-wire probe. The maximum axial distance between the wire and the orifice discharge plane was no more than $z = 3$ mm. The orifice discharge plane may actually have changed very slightly over one cycle due to slight out-of-plane motion of the false folds caused by rubber material deformation induced by the forcing action of the actuating rods. The approximate hot-wire discharge flow velocity measurement locations are shown in Fig. 4. In fact, however, the point of closure was not distributed along a perfectly straight line along the y direction. The abscissa of the point of closure was measured for each one of the seven latitudinal locations indicated in the figure. The probe was then traversed over exactly 1 mm on each side of that point. Despite this measure, other slight misalignments and imperfections occurred, which made the measured flow velocity profiles appear to be asymmetric in some cases, as discussed further below.

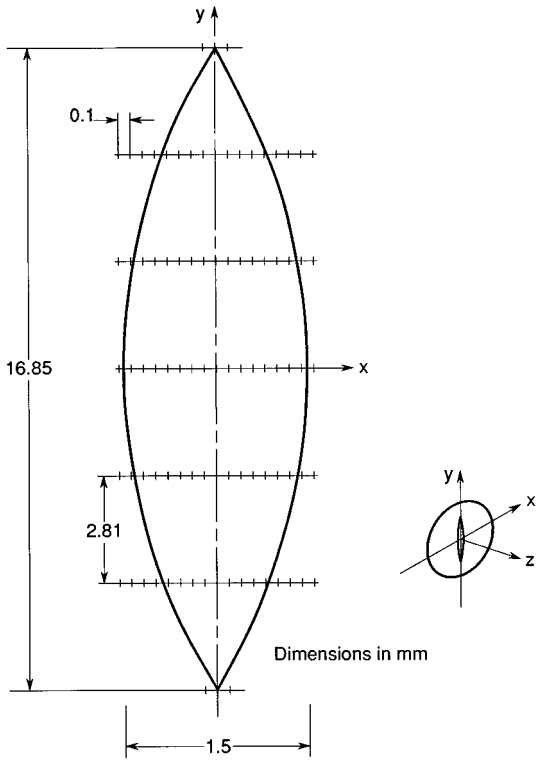


FIG. 4. Orifice dimensions and flow velocity measurement locations.

To obtain flow velocity, the hot-wire signals were processed using the following equation:

$$U = ak[(V^2 - esq)^{1/an}], \tag{6}$$

where V is the output voltage, with $esq = 0.501$, $an = 0.47$, and $ak = 24.576$. A TSI gas probe calibrator (model 1125) was used to calibrate the probe. The calibration was also frequently monitored *in situ*, using the glottal orifice itself as the calibrating apparatus. A static pressure rise of 1 mm Hg (13.6 mm water), corresponding to a flow velocity $U_{0c} = 20$ m/s from Eq. (1), always yielded an output voltage $V = 1.187$ V (dc).

The acoustic pressure was measured using a 6-mm-diam condenser microphone (Brüel & Kjaer type 4136) located at $z = -6.16$ cm, upstream of the orifice plane. Data were acquired, stored, and postprocessed using a Concurrent 6300 data acquisition computer in conjunction with Stanford In-

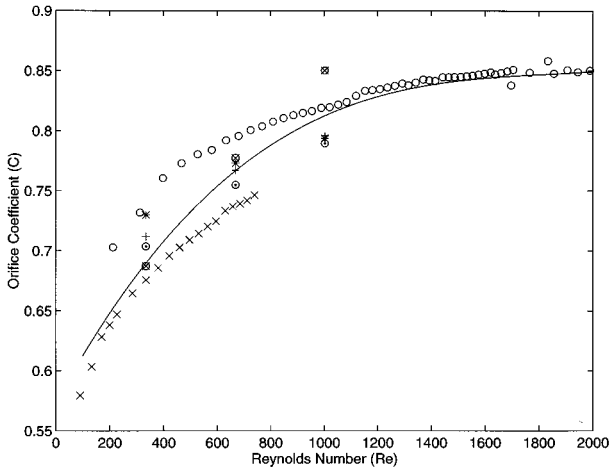


FIG. 5. Orifice coefficient versus Reynolds number for different orifice areas. o:$A_g=15.8$ mm^2, x:$A_g=8.4$ mm^2, *:$A_g=15.0$ mm^2, +:$A_g=11.5$ mm^2, ©:$A_g=7.5$ mm^2, ⊗:$A_g=4.0$ mm^2; ———: Eq. (8).

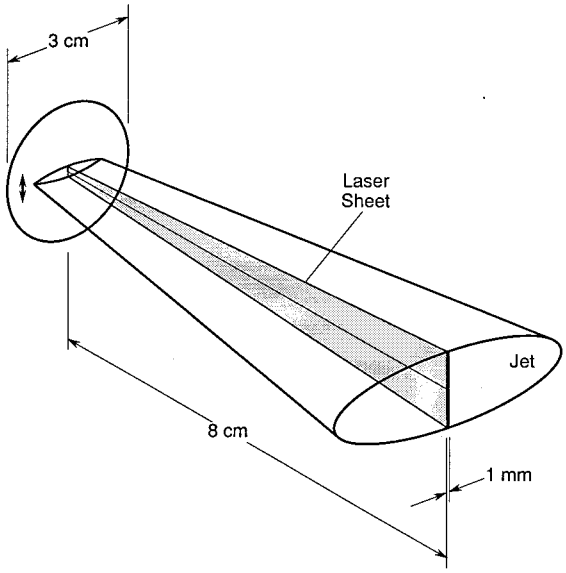


FIG. 6. Sketch showing the illuminated region of the jet for flow visualization.

struments SR-650 antialias analog filters and preamplifiers. The sampling frequency was 8192 Hz, which corresponded to 69 samples per cycle for a jet pulsation frequency of $\Omega=120$ Hz, with higher resolution at lower frequencies.

III. RESULTS

A. Static pressure-flow relationships

The static orifice flow resistance, $\Delta p_0/Q_0$, was first measured for different orifice areas A_g and approach Reynolds numbers Re (based on tube diameter and tube flow velocity),

$$\mathrm{Re}=\frac{Q_0 D_t}{A_t \nu}, \tag{7}$$

where D_t is the tube inside diameter and ν is the kinematic viscosity (m^2/s) Bernoulli's obstruction theory, Eq. (3), was then used to determine the orifice discharge coefficient C. The results are shown in Fig. 5. The coefficient was found to depend weakly on the area. The following approximate empirical relationship between orifice coefficient and Reynolds number was obtained, neglecting the influence of orifice area:

$$C(\mathrm{Re})=A(\mathrm{Re})^3+B(\mathrm{Re})^2+C(\mathrm{Re})+D, \tag{8}$$

with $A=3.89\times10^{-11}$, $B=-2.18\times10^{-7}$, $C=4.19\times10^{-4}$, and $D=5.75\times10^{-1}$. In the present case, $A_t=507$ mm^2, $\nu=1.5\times10^{-5}$ m^2 s^{-1}, and $D_t=2.54$ cm, which yields the simple relationship $\mathrm{Re}=3.343Q_0$, with the flow rate, Q_0,

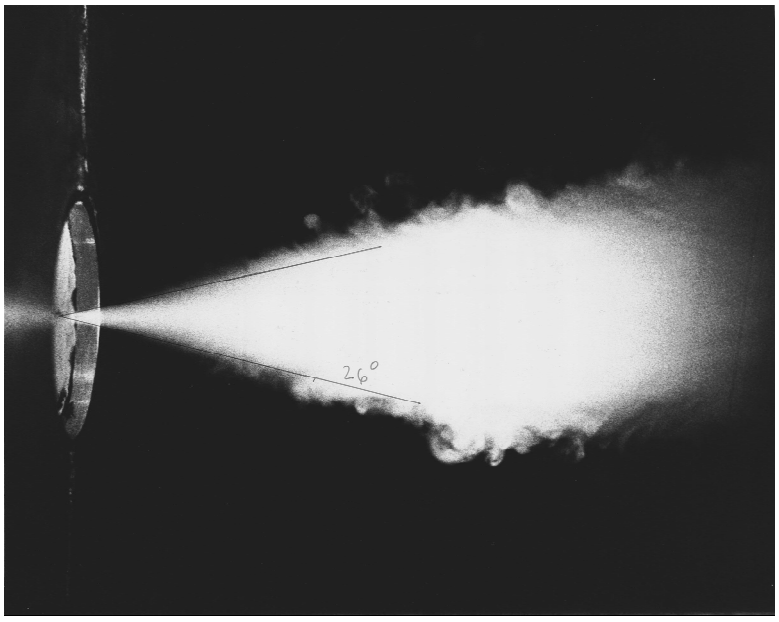


FIG. 7. Picture of the fully established pulsating jet. $p_{0t}=9$ cm water, $\Omega=60$ Hz.

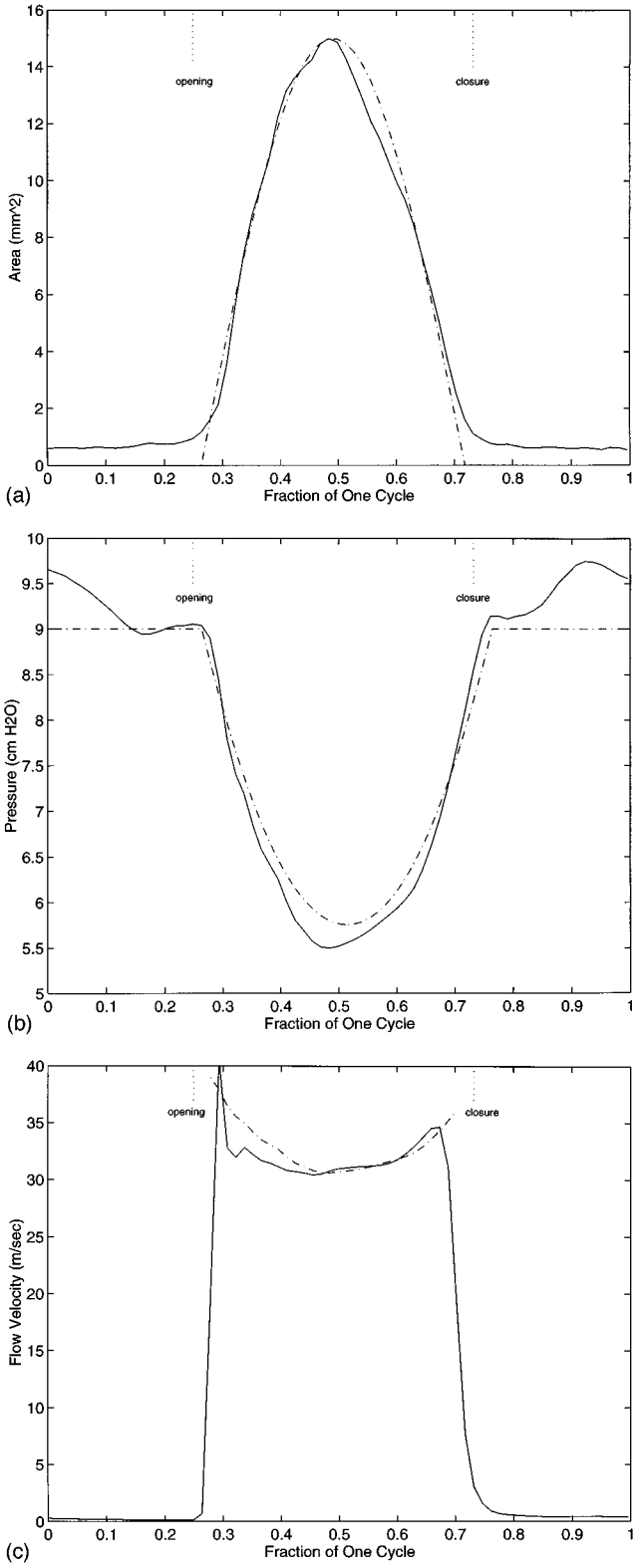


FIG. 8. Time history of orifice and flow variables for one period of oscillation. $\Delta p_{av}=8$ cm water, $\Omega=120$ Hz. (a) orifice area, ———: measured, —·—·—: Eq. (9); (b) pressure head, ———: measured, —·—·—: Eq. (10); (c) flow velocity in the center of the orifice, ———: measured, —·—·—: Eq. (1).

expressed in ml/s. These experimental values of the orifice coefficient, and Eq. (8), are in good general agreement with data from Scherer (1981) obtained using static models of the laryngeal airway.

B. Flow visualization

Flow visualization experiments were performed in order to detect possible changes in the flow characteristics of steady versus pulsating open jets. Cigarette smoke injected upstream of the orifice was used in conjunction with a laser sheet illuminating a cross section of the jet, as shown schematically in Fig. 6. Pictures were taken using a 35-mm camera (Canon). For this particular case, the mean transglottal pressure was 8 cm water, and the frequency was 60 Hz. No attempt was made to synchronize the pictures with the orifice motion.

A typical picture of the pulsating jet is shown in Fig. 7. The main characteristics of the pulsating jet were found to be qualitatively very similar to those of a comparable steady jet. The mean discharge flow angle, i.e., the angle between the shear layers and the normal to the plane of the glottis, was seemingly unaffected by the flow pulsation. The value of 13° is in good agreement with discharge angles reported in the literature for low Mach number rectangular open jets (Schlichting, 1979). Note that the thickness of the aluminum plate resting against the orifice plate, which is clearly seem in the picture, was 6 mm. This gives an indication of the size of the jet plume. Attempts were made to resolve the early stage of the developing jet immediately after opening of the orifice by taking a sequence of pictures at high speed, using a framing camera (Beckman & Whitley model 350). The illumination provided by the laser (0.7 W), however, was not sufficient for the film to be impressed. Pictures of steady jets and pulsating jets at a frequency of 120 Hz were also taken. The jet plume appeared to be similar in all cases, although the method is fairly crude and precludes accurate comparisons.

C. Glottal area

The orifice area was measured using a photoelectric sensor located in the subglottal tube (see Fig. 2). The orifice was illuminated with a strong light source. The optical sensor captured the light radiated through the orifice inside the subglottal tube, producing an electric signal proportional to the area of the orifice. The sensor was calibrated by measuring the area of the orifice at rest following a different approach. A pointer located very close to the orifice, between the orifice plate and the light source, was traversed using the two-dimensional micrometric hot-wire probe traverse. The shadow of the pointer on the orifice plate was used to accurately measure the position of the contour of the orifice (The location of the point of closure was also measured in a similar fashion, which helped define the hot-wire sensor locations reported below). The glottal area was then calculated by numerical integration.

The area waveform for one glottal cycle, shown in Fig. 8(a), was approximately a sinusoid with a duty factor of one-half. The orifice area can be approximated by the equation

$$A_g(t)=A_{g\,\max}\sin\omega t, \quad \omega t_0\leqslant\omega t\leqslant\omega t_0+\pi,$$
$$=0, \quad \omega t_0+\pi\leqslant\omega t\leqslant\omega t_0+2\pi, \tag{9}$$

where $A_{g\,\max}=15\ \mathrm{mm}^2$ and t_0 is the opening time of the orifice (the beginning of the duty cycle). This area waveform

is typical of speech production (Gauffin *et al.*, 1981). It must be noted again that the maximum area was not always absolutely constant, and the orifice geometry not perfectly regular as discussed in Sec. II. In one series of experiments, however, the photoelectric signals were found to be independent of frequency (between 10 and 150 Hz). This indicated that the slight orifice deformations were induced by the forcing rods rather than inertial stresses. For the dynamic flow characteristics reported below, the maximum area (which was measured before, during, and after the experiment) was 15 mm^2 (within 3%).

D. Pressure head

The fluctuating pressure head, $\Delta p(t)$, across the orifice was calculated by summing the time-averaged static pressure read by the static pressure transducer, and the fluctuating pressure signal from the upstream microphone. The constant discharge pressure assumption was verified by probing the sound pressure in the vicinity of the orifice, outside the tube and away from the air stream. The supraglottal pressure fluctuations were lower than 0.1 cm water, and were primarily caused by the background mechanical noise of the apparatus. Since the radiation impedance of an orifice is much smaller in free space than in a duct, the open jet did not radiate much sound outside the tube compared with the sound waves radiated inside the subglottal tube. These operating conditions are different from actual speech, where significant pressure fluctuations occur in the vocal tract. Removing the vocal tract has an effect comparable to that of short-circuiting the load impedance in the well-known electroacoustic equivalent model of the vocal apparatus (Flanagan, 1965). Since the subglottal impedance is known and purely resistive, the only unknown in the equivalent circuit is then the glottal resistance (usually a variable resistance, with any inductive effect neglected). This approach was purposefully adopted for this study because it eliminated the complications associated with the presence of a reactive acoustic load. The effects related to the presence of a pseudo-vocal tract will be the object of future investigations.

Figure 8(b) shows a typical pressure waveform measured by the pressure sensor in the subglottal tube. A time delay of a little more than one time step occurred between glottal opening and the arrival of the pressure disturbance at the microphone location, which is consistent with sonic propagation ($z_{\mathrm{mic}}/c_0 = 0.0616/344 = 0.18$ ms or about 2% of the period in this case). The magnitude of the subglottal pressure fluctuations, about 3 cm of water, corresponded to 35% of the mean static pressure drop, or a sound level of 143 dB (*re:* 20 μPa).

Shown in Fig. 8(b) is the pressure predicted by the analytical model of Sec. II for a blocked pressure $p_{0t} = 9$ cm water, a maximum orifice area of $A_{g\,\max} = 15\ mm^2$, and a constant orifice coefficient of $C = 0.85$. The following quadratic equation for the pressure, obtained by combining Eqs. (3) and (5), was solved analytically:

$$p'(t)^2 - 2\rho_0 c_0^2 C^2 \left[\frac{A_g(t)}{A_t}\right]^2 [p'(t) + \Delta p_0] = 0. \qquad (10)$$

The agreement between experimental data and theory is very good, which indicates the validity of both the quasi-steady approximation and the source model. Discrepancies were attributed to sound wave reflections within the tube due to imperfect anechoic termination, a common problem for small ducts at relatively low frequencies. The subglottal duct input impedance, measured using a two-sensor method, indicated the presence of a standing wave, implying a reflective termination.

E. Velocity

The instantaneous jet flow velocity profile was measured using hot-wire anemometry. The measurement locations, distributed on a plane located at approximately $z = 3$ mm away from the orifice discharge, are shown in Fig. 4. The methodology was described in Sec. II. Only the first portion of the velocity records (which were 1024 bytes long, corresponding to 14 complete cycles at 120 Hz) is shown in the following. The signals were acquired in synchronization with the motion of the glottis. A tachometer signal was used to trigger the data acquisition computer clocks.

Figure 8(c) shows the waveform of the flow velocity at the center of the orifice. The mean (time-averaged) pressure head was 8 cm water, and the frequency was 120 Hz. The instantaneous velocity waveform was smooth, and showed a variation of less than 1% from one cycle to the next. It did not exhibit a random component, which indicated a laminar flow. It is therefore likely that the probe was located within the potential core of the jet. Previous studies using excised canine larynges reported significant turbulence in the velocity waveform (Berke *et al.*, 1989; Alipour and Scherer, 1995). However, these measurements were made further downstream from the orifice, because of physical interference between dog larynx cartilage and the probe traversing system. The instantaneous velocity measurements were therefore probably made in the mixing region and the fully developed region where the flow may have been turbulent. Evidence of turbulence in the fully developed part of the jet is clearly suggested by Fig. 7. No hot-wire measurements, however, were made to confirm this hypothesis.

Precisely at the moment where the orifice opened, the instantaneous discharge flow velocity increased very rapidly until it reached a value around 40 m/s, or approximately Mach 0.12. The velocity then dropped to around 30 m/s, before increasing again slightly and regularly. The velocity then dropped sharply to a very small value just after closure.

The instantaneous discharge flow velocity predicted by Bernoulli's equation (assuming the flow is quasi-steady) was calculated using Eq. (1) and measured dynamic pressure head data [shown in Fig. 8(b)]. The resulting flow velocity, shown in Fig. 8(c), is in good agreement with the measured velocity waveform for most of the duty cycle. This observation confirms the validity of the quasi-steady assumption, within experimental accuracy, for a significant portion of the duty cycle. At orifice opening, however, the sudden jump observed in the measured flow velocity signal is not predicted by Bernoulli's equation. The velocity jump is followed by an overshoot, before the velocity reaches the predicted value.

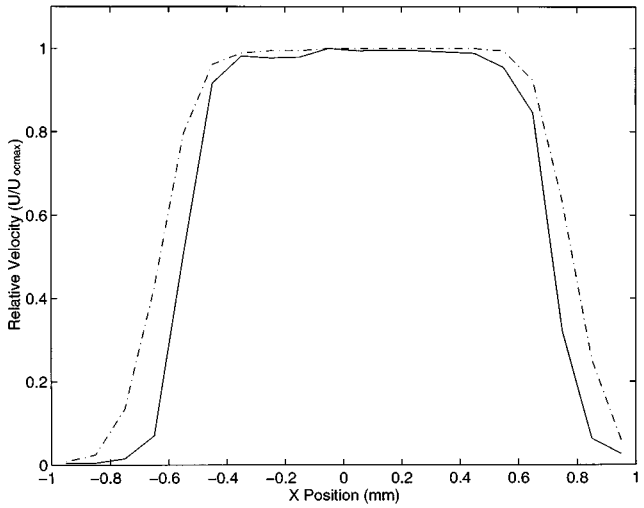


FIG. 9. Comparison between the relative velocity profiles measured along the x axis (at $y=0$) for steady and unsteady jets. ———: instantaneous, measured for $p_{0t}=9$ cm water and $\Omega=120$ Hz; —·—·—: measured for a comparable, steady flow.

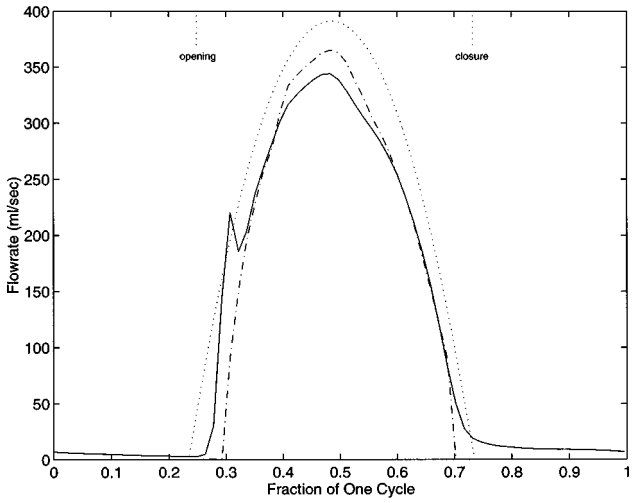


FIG. 10. Flow rate for one period of oscillation. ———: measured, obtained by numerical integration of the velocity profile; ·····: calculated from Eqs. (5), (8), and (9) with $C=0.85$; —·—·—: calculated from Eqs. (3) and (8), using experimental data for pressure and area.

It may be worth mentioning that the orifice did not close uniformly, but with a slight, variable time-delay along the y direction, which might explain the apparent delay between velocity decay and closure.

F. Flow rate

Instantaneous velocity profiles across the glottis were measured at different latitudinal positions y as a function of time. Figure 9 shows the instantaneous velocity profile along the x axis at the instant where the orifice area was maximum. This velocity profile was compared with that of a static jet measured using the same methodology, but for a pressure head of 5.73 cm H_2O and with a fixed orifice geometry. The relative velocity profiles, shown in Fig. 9, are in good general agreement. The pulsating jet was slightly skewed with respect to the axial direction, but the shear layers thickness and the overall velocity profiles were similar. The difference in the width of the jets was probably because the maximum area was 17.8 mm^2 for the static measurement, but only 15 mm^2 for the dynamic measurement (the area changes were caused by slippage of mechanical components, as mentioned in Sec. II). Several adjustments were made to the apparatus between the static measurements and the dynamic measurements, which were performed at different times. There is no sign of additional flow turbulence or flow nonuniformity related to the pulsating character of the jet. This was found to be the case not only at the instant where the area is maximum, but for the entire glottal cycle. The instantaneous velocity profiles were always smooth and similar to those shown in Fig. 9. Note, however, that single-wire hot-wire probes do not allow the resolution of two velocity components, therefore the measured velocity amplitudes within the shear layer could be inaccurate.

The instantaneous flow rate, Q, was obtained at every time step by integrating the instantaneous velocity profiles over the surface of the orifice numerically, using a trapezoidal rule. The result is shown in Fig. 10. Although the flow rate waveform is (as expected) very close to that of the orifice area, Fig. 8(a), a burst at the beginning of the duty cycle similar to that in the velocity signals can be seen. The flow burst was detected for most probe locations in the vicinity of the orifice, including off-centered locations. Note again, however, that single hot-wire probes are not sensitive to flow direction. The flow direction is not clear for the glottal opening stage since the detailed flow structure associated with the incipience of the flow burst is not known. Therefore it is possible that negative flow velocities may have been improperly accounted for in the flow rate calculation. For the majority of the cycle, however, there is no ambiguity in flow direction since the jet appears to be nearly quasi-steady. The time-averaged volume flux calculated from the space-averaged velocity signals, 112 ml/s, was in excellent agreement with the mean flow rate read by the mass flow meter, 108 ml/s, which supports the accuracy of the measured flow velocities.

The flow rate calculated from the theoretical model of Sec. I [with p_{0t}, C, and A_g prescribed as for the pressure calculation in Fig. 8(b)] is shown in Fig. 10. For comparison, the flow rate, Q, calculated from Eq. (3) using measured pressure and area data, as well as Eq. (8) for the orifice coefficient is also shown in the same figure. Good agreement between the experimental results and the theoretical models was again found for most of the duty cycle, after the jet is fully established. This result therefore supports the validity of the quasi-steady assumption for an open jet through a converging orifice with a time-varying area. The early phase of flow development within the orifice is impulsive in nature, and does not obey the quasi-steady model. This transient phenomenon does not seem to contribute to the sound pressure wave radiated in the subglottal tube, shown in Fig. 8(b). According to Eq. (5), a jump in flow rate at glottal opening should result in a comparably abrupt drop in the pressure waveform. Instead, the pressure drop is comparatively gradual with no overshoot as for the case of the flow rate waveform.

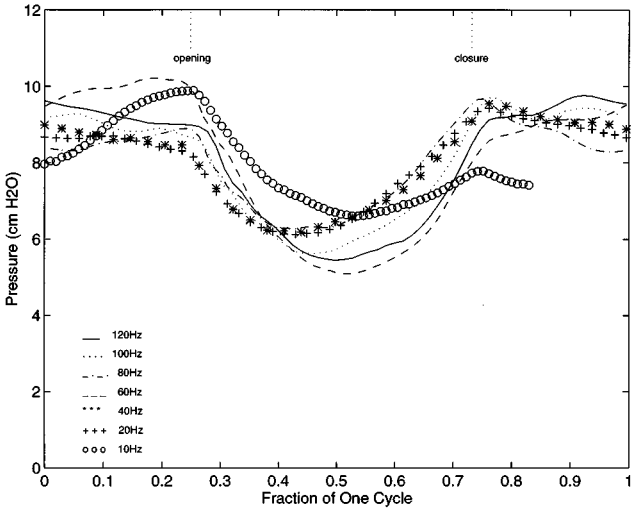


FIG. 11. Effects of frequency on the pressure pulsations in the tube.

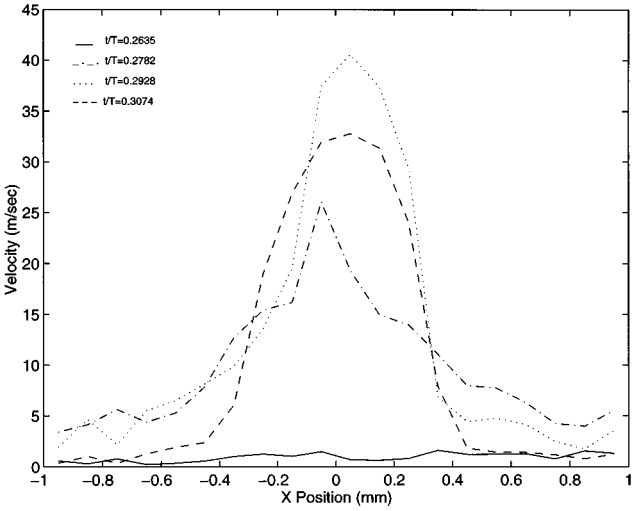


FIG. 13. Velocity profile along the x direction ($y=0$) at four successive times, immediately after orifice opening.

IV. FURTHER RESULTS AND DISCUSSION

A. The effects of frequency

The pressure head waveforms for frequencies of 10, 20, 40, 60, 80, 100, and 120 Hz are shown in Fig. 11. The Strouhal number based on the wetted perimeter of the glottal orifice and jet discharge flow velocity, $St=f\sqrt{A_{g\,\max}}/U_{0c\max}$, varied within the range $0\leqslant St\leqslant 0.02$. The maximum area drifted only slightly (less than 2%) over the duration of the experiments. The pressure head curves are in good general agreement. The significant discrepancies for certain frequencies were attributed to the effects of wave reflections from the tube end due to imperfectly anechoic termination, as discussed earlier. In general, the time history of the flow was independent of frequency (or Strouhal number). For example, Fig. 12 shows a typical plot of the velocity at 10 Hz for the same mean pressure head, with the probe located off-center ($y=0$, $x=0.3$ mm, same z location). Once again, the velocity predicted from Eq. (1) using measured pressure data is shown in the figure. The flow burst is seen at the start of the duty cycle. The velocity quickly reaches the quasi-steady value after the developed jet wets the hot-wire probe (the jet area increases with time). The jet velocity during the fully developed part of the duty cycle was equally well predicted by Bernoulli's equation for the entire frequency range investigated.

B. Glottal opening stage

Figure 13 shows a sequence of velocity profiles measured at regular time intervals, immediately after glottal opening. For this case, data were sampled at 8190 Hz and the glottal oscillation frequency was 120 Hz. Therefore each time interval corresponded to about 2% of one glottal cycle. Close examination of the evolution of the flow distribution after glottal opening shows that, immediately after opening, the velocity distribution is widespread across the glottal

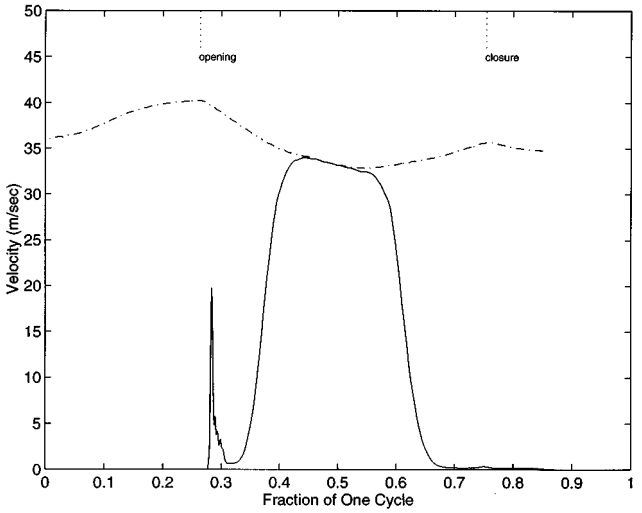


FIG. 12. Velocity for one period of flow oscillations, measured off-center at $x=0.3$ mm and $y=0$. $\Omega=10$ Hz, $\Delta p_{\mathrm{av}}=8$ cm H_2O. ———: measured, —·—·—: Eq. (1).

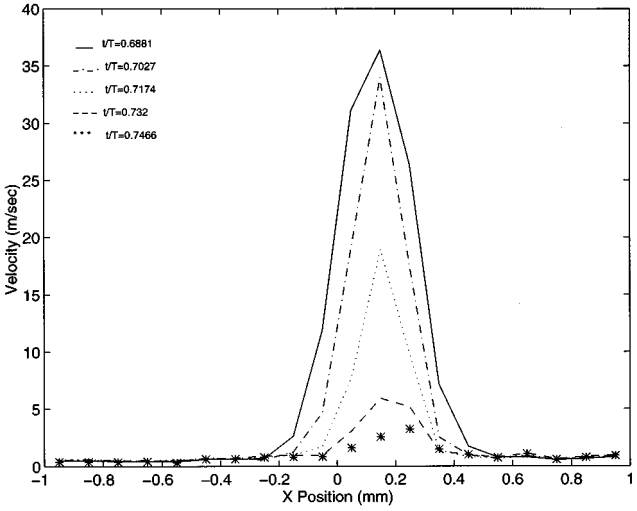


FIG. 14. Velocity profile along the x direction ($y=0$) at four successive times, immediately after orifice closure.

plane. The burst of flow rapidly decays to be replaced by a jet flow, whose characteristics appear to be similar to a steady jet. Note again that inaccuracies in the flow velocity amplitude and direction for the transient phenomenon are to be expected due to the limitations of the hot-wire anemometer. For comparison, the evolution of the velocity profile during closure, shown in Fig. 14, is regular. The apparent asymmetries in the flow velocity profiles are due in part to the spatial resolution of the measurements, but also to the slight imperfections in the orifice geometry outlined in Sec. II. There is no other possible reason, since turbulent quasi-steady jets through perfectly symmetrical orifices should be symmetric.

An interesting discussion of developing, impulsively started jets (and their impingement on a fixed obstacle) is provided by Verge *et al.* (1994) for the case of an organ flue pipe. Detailed flow visualization and pressure measurements were performed using a fixed orifice. A sliding gate was installed between the flow supply and the duct upstream of the orifice. The device was rapidly opened, creating an impulsively started flow within the orifice. The velocity of the flow structure associated with jet formation was measured. Different flow models, based on transient acoustic equations and the unsteady Bernoulli equation, were suggested to predict the flow behavior. None of the models, however, provided very accurate predictions. In addition, flow acceleration was limited by the time required to open the sliding valve, which was around 0.5 ms. The flow structure of the resulting developing, incompressible viscous jet involved the formation of a vortex ring rolled up around the contour of the orifice discharged by the forming laminar shear layer. As time progressed, the vortex ring grew, and was entrained away from the wall by the forming jet upstream. At that point, the flow structure of the developing incompressible jet resembled a mushroom. Eventually, the vortex ring gradually disappeared to be replaced by a fully established jet. No evidence of a flow burst associated with the formation of the jet was reported.

The jet formation stage for an incompressible, impulsively started jet through a fixed orifice was investigated numerically by Shimizu *et al.* (1986) and Shimizu and Kokoshima (1990). A finite difference method allowed the calculation of the jet flow velocity as a function of time. It was found that the velocity in the center core of the developing jet temporarily reached a value that was higher than that for the subsequent, developed phase of the jet. The formation of the vortex ring was found to occur over a time period similar to that reported in Verge *et al.* (1994). The reported velocity increase, however, seemed to occur after a time delay that was long compared with that observed in the case of the present study, where the disturbance appears to be propagating very fast, at a speed closer to sonic than to jet velocity.

It seems therefore that forced flow through suddenly opened orifices may have different characteristics than impulsively started flows through static orifices. Although no conclusive explanations or models were found to explain the sudden flow acceleration associated with the release of the folds, the following hypothesis was formulated. The sudden rise in flow velocity could be an acoustic ''near-field'' effect, involving a combination of inertial, compression, and viscous stresses. At time $t=t_0^-$, just before orifice opening, the flow on each side of the closed orifice is at rest. A discrete pressure gradient (of about 10 cm water) is present between the two sides of the orifice [see Fig. 8(b)]. At time t_0^+, when the orifice opens, a very steep pressure gradient exists within the airway created by the release of the false folds (the pressure distribution is analogous to a step function). This could result in a very large flow acceleration, leading to a sudden expansion of the higher pressure, higher density fluid within the airway towards the lower pressure region where the hot-wire probe was located. A small amount of air mass could therefore have been expelled out of the flow passage immediately after fold release, before any motion of the fluid within the tube upstream of the orifice could take place. The sudden flow expansion may have established very quickly a constant pressure within the glottal passage, which was then followed by the establishment of flow from the tube through the orifice, which eventually resulted in the formation of the jet flow. According to this hypothesis, flow compressibility must therefore be taken into account to explain the initial flow ''burst,'' which involves, perhaps, something similar to a very weak shock wave. This phenomenon is not predicted by Bernoulli's equation. The fact that the strong velocity surge recorded by the hot-wire probe did not affect the pressure in the tube (the microphone was located roughly 6 cm, or more than two duct diameters upstream from the orifice) may be related to the fact that the phenomenon is a ''near-field'' effect. A pressure sensor mounted within the orifice wall may have registered a sudden pressure drop after glottal opening. Because of viscous stresses, the fluid mass rapidly expelled from the orifice may still roll up into vortex rings, as for incompressible starting jets. Work is needed to provide a more detailed description of the phenomenon.

This abrupt flow acceleration may be more a property of the mechanical model than of real speech. Detailed action of the driven rubber model may not reflect real glottal behavior at the precise moments of closure and release. Most of the time, real speech shows little or no evidence of abrupt glottal release; a more gradual release may reduce the burst. Also, much of the time the real glottis does not close completely; there remains a small opening between the arytenoids. In that case, a burst would probably not occur since the pressure gradient responsible for the flow acceleration would be much smaller immediately after release.

C. Importance of accounting for flow acceleration

For unsteady potential incompressible flows, the effects of flow acceleration may be accounted using the so-called ''unsteady Bernoulli's equation'' [see, for example, Wylie and Streeter (1978)]. Integration of Euler's equation for unsteady, incompressible, irrotational flow, neglecting gravity, yields

$$\frac{\partial \phi}{\partial t}+\frac{p}{\rho}+\frac{1}{2}|\mathbf{u}|^2=f(t), \tag{11}$$

where the velocity $u = \nabla \phi$ can be obtained from the velocity potential ϕ, and f is one arbitrary function of time. The flow velocity within the jet center core was calculated from Eq. (11), as well as measured pressure and area data. The influence of flow acceleration [the first term in Eq. (11)] was estimated by comparing the predictions obtained using both the steady and the unsteady versions of Bernoulli's equation. Besides at opening and closure, however, the effects of flow acceleration were not significant, i.e., the velocity predictions from Eqs. (3) and (11) were almost identical. This was probably because the velocity remained nearly constant for most of the duty cycle in this case. For release and closure, the unsteady Bernoulli equation predicted an exponential buildup and decay [similarly to the step function response of a resistive and capacitive (RC) electrical circuit]. These trends, however, did not qualitatively agree very well with the observed velocity waveform at glottal release and closure [see Fig. 8(c)]. At opening, the actual flow velocity increased suddenly rather than exponentially. At closure, the observed decay rate was much more pronounced than that predicted by the unsteady Bernoulli equation. These observations suggest that viscous effects and acoustic near-field effects must be accounted for in order to accurately model the details of the flow associated with vocal fold release, which will be the object of future work.

V. CONCLUSIONS

Experiments were performed in order to verify the legitimacy of the quasi-steady assumption often made when modeling voice production in articulatory speech synthesis and speech recognition models. Detailed flow measurements of a pulsating open jet through a modulated orifice with a time-varying area indicated that the assumption was valid, and yields accurate pressure and flow rate predictions during the portion of the glottal cycle where the jet is fully established. As foreseen by Flanagan (1965), the assumption breaks down during the jet formation stage, which corresponded to about one-fifth of the duty cycle. A flow burst occurred immediately after opening of the orifice. The associated flow structure may involve fluid compressibility and the formation of an elongated vortex ring entrained by the main core of the forming jet. The phenomenon caused the flow velocity to impulsively reach a peak immediately after opening. This transient, however, did not affect the signature of the pressure waves generated at the orifice and propagated upstream along the ''subglottal'' tube. These pressure waves were otherwise predicted very well by a simple one-dimensional monopole model. Further work is under way to investigate a confined pulsating jet flow, in the presence of acoustical loading.

ACKNOWLEDGMENTS

The essential part of this project was conducted while the first author was a postdoctoral member of technical staff at AT&T Bell Laboratories (now Bell Laboratories) in Murray Hill, NJ. The authors are very grateful to Dr. Peter Simpkins for his assistance with the flow visualization experiments, and many helpful discussions. Thanks are also expressed to Mr. Daniel Quinlan for his assistance and support, and to Dr. Richard McGowan, Dr. Ingo Titze, Dr. Mohan Sondhi, Dr. Juergen Schroeder, and Dr. Ronald Scherer for their helpful comments and discussions.

Alipour, F., and Scherer, R. C. (**1995**). ''Pulsatile airflow during phonation: An excised larynx model,'' J. Acoust. Soc. Am. **97,** 1241–1248.

Alipour, F., Fan, C., and Scherer, R. C. (**1996**). ''A numerical simulation of laryngeal flow in a forced-oscillation glottal model,'' Comput. Speech Lang. **10,** 75–93.

Allen, D. R., and Strong, W. J. (**1985**). ''A model for the synthesis of natural sounding vowels,'' J. Acoust. Soc. Am. **78,** 58–69.

Anathapadmanabha, T. V., and Fant, G. (**1982**). ''Calculation of the true glottal flow and its components,'' Speech Transmiss. Lab. Q. Prog. Stat. Rep. 1/1982, Stockholm, 1–30.

Beranek, L. L. (**1986**). *Acoustics* (American Institute of Physics, New York).

Berke, G. S., Moore, M. D., Monkewitz, P. A., Hanson, D. G., and Gerratt, B. R. (**1989**). ''A preliminary study of particle velocity during phonation in an *in vivo* canine model,'' J. Voice **3**, 306–313.

Berke, G. S., Green, D. C., Smith, M. E., Arnstein, D. P., Honrubia, V., Natividad, M., and Conrad, W. A. (**1991**). ''Experimental evidence in the *in vivo* canine for the collapsible tube model of phonation,'' J. Acoust. Soc. Am. **89,** 1358–1363.

Binh, N., and Gauffin, J. (**1983**). ''Aerodynamic Measurements in an Enlarged Static Laryngeal Model,'' R. Inst. Tech. Stockholm, STL-QPSR **2–3**, 36–60.

Bousfield, D. W., Stockel, I. H., and Nanivadekar, C. K. (**1990**). ''The breakup of viscous jets with large velocity modulations,'' J. Fluid Mech. **218,** 601–617.

Cranen, B. (**1987**). ''The acoustic impedance of the glottis,'' Doctoral Dissertation, Katholieke Universiteit, Nijmegen.

Davies, H. G., and Fflowcs Williams, J. E. (**1968**). ''Aerodynamic sound generation in a pipe,'' J. Fluid Mech. **32,** 765–778.

Diebold, B., Delouche, A., Dumée, J. P., Guglielmi, J. P., Delouche, P., and Péronneau, P. (**1990**). ''*In vitro* analysis of a model of intracardiac jet: analysis of the central core of axisymmetric jets,'' J. Biomech. **23,** 35–44.

Dowling, A. P., and Ffowcs Williams, J. E. (**1983**). *Sound and Sources of Sound* (Ellis Horwood, England).

Fant, G. (**1982**). ''Preliminaries to Analysis of the Human Voice Source,'' R. Inst. Tech. Stockholm, STL-QPSR **4**, 1–27.

Flanagan, J. L. (**1965**). *Speech Analysis, Synthesis, and Perception* (Springer-Verlag, Berlin).

Flanagan, J. L., and Landgraf, L. L. (**1968**). ''Self-oscillating source for vocal-tract synthesizers,'' IEEE Trans. Audio Electroacoust. **AU-16,** 57–64.

Fletcher, N. H. (**1979**). ''Air flow and sound generation in musical wind instruments,'' Annu. Rev. Fluid Mech. **11,** 123–146.

Gauffin, J., Binh, N., Ananthapadmanabha, T. V., and Fant, G. (**1981**). ''Glottal geometry and volume velocity waveform,'' Proc. Res. Conf. Voice Physiology, Madison, WI.

Guo, C. G., and Scherer, R. C. (**1993**). ''Finite element simulation of glottal flow and pressure,'' J. Acoust. Soc. Am. **94,** 688–700.

Gupta, S. K., and Schroeter, J. (**1993**). ''Pitch-synchronous frame-by-frame and segment-based articulatory analysis by synthesis,'' J. Acoust. Soc. Am. **94,** 2517–2531.

Hirschberg, A., van de Laar, R. W. A., Marrou-Maurières, J. P., and Wijnands, A. P. J. (**1990**). ''A quasi-stationary model of air flow in the reed channel of single-reed woodwind instruments,'' Acustica **70,** 146–154.

Iguchi, M., Yamazaki, H., Yamada, E., and Morita, Z. (**1990**). ''Velocity and turbulence intensity in a pulsating jet through a sudden expansion,'' Trans. Jpn. Soc. Mech. Eng. **56**, 1659–1664.

Ikeda, T., Matsuzaki, Y., and Sasaki, T. (**1994**). ''Separated flow in a channel with an oscillating constriction (numerical analysis and experiment),'' Trans. Jpn. Soc. Mech. Eng. **60**, 750–757.

Ingård, U., and Labate, S. (**1950**). ''Acoustic circulation effects and nonlinear impedance of orifices,'' J. Acoust. Soc. Am. **22,** 211–218.

Ingård, U. (**1953**). ''On the theory and design of acoustic resonators,'' J. Acoust. Soc. Am. **25,** 1037–1061.

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,** 1233–1268.

Ishizaka, K., and Matsudaira, M. (**1972**). ''Fluid mechanical considerations of vocal fold vibration,'' SCRL (Speech Communication Research Laboratory) Monograph No. 8, Santa Barbara, CA.

Kaiser, J. (**1983**). ''Some observations on vocal tract operation from a fluid flow point of view,'' in *Vocal Fold Physiology: Biomechanics, Acoustics, and Phonatory Control*, edited by I. R. Titze and R. H. Scherer (Denver Center for the Performing Arts, Denver), pp. 358–386.

Kathuriya, M. L., and Munjal, M. L. (**1979**). ''Experimental evaluation of the aeroacoustic characteristics of a source of pulsating gas flow,'' J. Acoust. Soc. Am. **65,** 240–248.

Koizumi, T. K., Taniguchi, S., and Hiromitsu, S. (**1985**). ''Glottal source-vocal tract interaction,'' J. Acoust. Soc. Am. **78,** 1541–1547.

Kozenko, G. D., Sapozhnikov, B. G., and Usenko, Y. A. (**1989**). ''Experimental investigation of the velocity of alternating jets issuing from a horizonal annular cavity with a vibrating inner cylindrical surface,'' Izv. Akad. Nauk SSSR, Mekh. Zh.-Gaza, **4**, 178–180 (in Russian).

Laine, U., and Karjalainen, M. (**1986**). ''Measurements of the effects of glottal opening and flow on the glottal impedance,'' Proc. ICASSP 86, Tokyo, 1621–1624.

McGowan, R. S. (**1988**). ''An aeroacoustic approach to phonation,'' J. Acoust. Soc. Am. **83,** 696–704.

McGowan, R. S. (**1992**). ''Tongue-tip trills and vocal tract wall compliance,'' J. Acoust. Soc. Am. **91,** 2903–2910.

McGowan, R. S. (**1993**). ''The quasisteady approximation in speech production,'' J. Acoust. Soc. Am. **94,** 3011–3013.

Pedley, T. J. (**1983**). ''Wave phenomena in physiological flows,'' Appl. Math. **32**, 267–287.

Pedley, T. J., and Stephanoff, K. D. (**1985**). ''Flow along a channel with a time dependent indentation in one wall: The generation of vorticity waves,'' J. Fluid Mech. **160,** 337–367.

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.

Pelorson, X., Hirschberg, A., Wijnands, A. P. J., and Bailliet, H. M. A. (**1995**). ''Description of the flow through *in-vitro* models of the glottis during phonation,'' Acta Acust. **3**, 191–202.

Pelorson, X., Vescovi, C., Castelli, E., Hirschberg, A., Wijnands, A. P. J., and Bailliet, H. M. A. (**1996**). ''Description of the flow through *in-vitro* models of the glottis during phonation: application to voiced sound synthesis,'' Acta Acust. **82**, 358–361.

Pierce, A. D. (**1989**). *Acoustics: An Introduction to its Physical Principles and Applications* (American Institute of Physics, New York).

Rösler, S., and Strube, W. (**1989**). ''Measurement of the glottal impedance with a mechanical model,'' J. Acoust. Soc. Am. **86,** 1708–1716.

Scherer, R. C. (**1981**). ''Laryngeal fluid mechanics: steady flow considerations using static models,'' Ph.D. Thesis, University of Iowa, Iowa City.

Scherer, R. C. (**1991a**). ''Aerodynamics of voice production,'' in *Assessment of Speech and Voice Production: Research and Clinical Applications*, edited by J. A. Cooper (National Institute on Deafness and Other Communication Disorders, NIH).

Scherer, R. C. (**1991b**). ''Physiology of Phonation: A Review of Basic Mechanics,'' in *Phonosurgery: Assessment and Surgical Management of Voice Disorders*, edited by C. N. Ford and D. M. Bless (Raven, New York).

Scherer, R. C., Titze, I. R., and Curtis, J. F. (**1983**). ''Pressure-flow relationships in two models of the larynx having rectangular glottal shapes,'' J. Acoust. Soc. Am. **73,** 668–676.

Scherer, R. C., and Guo, C. G. (**1990**). ''Laryngeal modeling: Translaryngeal pressure for a model with many glottal shapes,'' *ICSLP Proceedings*, 1990 International Conference on Spoken Language, J. Acoust. Soc. Jpn. **1,** 3.1.1–3.1.4.

Schlichting, H. (**1979**). *Boundary-Layer Theory* (McGraw–Hill, New York).

Shadle, C. H., Elliott, S. J., and Nelson, P. A. (**1987**). ''Visualization of the air flowing through a dynamic model of the vocal folds,'' ISVR Technical Report 154.

Shimizu, A., Ishii, H., and Tsutomu, W. (**1986**). ''A numerical analysis of vortex growth in a bounded rectangular jet,'' Comput. Fluids **14**(4), 327–359.

Shimizu, A., and Kokoshima, Y. (**1990**). ''A numerical analysis of a two-dimensional impulsively started jet flow,'' Trans. Jpn. Soc. Mech. Eng. **56**, 690–695.

So, R. M. C., Ahmed, S. A., and Yu, M. H. (**1987**). ''The near field behavior of turbulent gas jets in a long confinement,'' Exp. Fluids **5,** 2–10.

Sondhi, M. M. (**1975**). ''Measurement of the glottal waveform,'' J. Acoust. Soc. Am. **57,** 228–232.

Teager, H., and Teager, S. (**1983**). ''Active fluid dynamic voice production models, or is there a unicorn in the garden?,'' in *Vocal Fold Physiology: Biomechanics, Acoustics, and Phonatory Control*, edited by I. R. Titze and R. H. Scherer (Denver Center for the Performing Arts, Denver), pp. 387–401.

Titze, I. R. (**1973**). ''The human vocal cords: A mathematical model. Part I,'' Phonetica **28**, 128–170.

Titze, I. R. (**1974**). ''The human vocal cords: A mathematical model. Part II,'' Phonetica **29**, 1–21.

Titze, I. R., and Talkin, D. T. (**1979**). ''A theoretical study of the effects of various laryngeal configurations on the acoustics of phonation,'' J. Acoust. Soc. Am. **66,** 60–74.

van den Berg, J. W., Zantema, J. T., and Doornenbal, P. (**1957**). ''On the air resistance and the Bernouilli effect of the human larynx,'' J. Acoust. Soc. Am. **29,** 626–631.

Verge, M. P., Fabre, B., Mahu, W. E. A., Hirschberg, A., van Hassel, R. R., and Wijnands, A. P. J. (**1994**). ''Jet formation and jet velocity fluctuations in a flue organ pipe,'' J. Acoust. Soc. Am. **95,** 1119–1132.

Wegel, R. L. (**1930**). ''Theory of vibration of the larynx,'' Bell Syst. Tech. J. **9,** 207–227.

Wylie, E. B., and Streeter, V. L. (**1978**). *Fluid Transients* (McGraw–Hill, New York).

White, F. M. (**1986**). *Fluid Mechanics* (McGraw–Hill, New York).

Perceptual compensation for coarticulation by Japanese quail (*Coturnix coturnix japonica*)

Andrew J. Lotto,[a] Keith R. Kluender, and Lori L. Holt
University of Wisconsin, Madison, Wisconsin 53706

(Received 18 April 1996; revised 19 March 1997; accepted 16 April 1997)

When members of a series of synthesized stop consonants varying in third-formant ($F3$) characteristics and varying perceptually from /da/ to /ga/ are preceded by /al/, human listeners report hearing more /ga/ syllables than when the members of the series are preceded by /ar/. It has been suggested that this shift in identification is the result of specialized processes that compensate for acoustic consequences of coarticulation. To test the species-specificity of this perceptual phenomenon, data were collected from nonhuman animals in a syllable ''labeling'' task. Four Japanese quail (*Coturnix coturnix japonica*) were trained to peck a key differentially to identify clear /da/ and /ga/ exemplars. After training, ambiguous members of a /da/–/ga/ series were presented in the context of /al/ and /ar/ syllables. Pecking performance demonstrated a shift which coincided with data from humans. These results suggest that processes underlying ''perceptual compensation for coarticulation'' are species-general. In addition, the pattern of response behavior expressed is rather common across perceptual systems. [S0001-4966(97)01608-1]

PACS numbers: 43.71.An, 43.71.Es, 43.80.Lb [WS]

INTRODUCTION

One of the wonders of speech communication is the remarkable symmetry between perception and production. Articulatory dynamics and constraints shape the resultant waveform, and in many cases, perceptual processes of the listener act in ways that appear to respect these constraints and dynamic properties. The beauty of this symmetry (or synergy) has motivated several comprehensive theories of speech communication. For example, in the revised Motor Theory of Liberman and Mattingly (1985), symmetry is said to arise because of the shared currency of speech perception and production: gestural representations. Auditorist theories, such as espoused by Diehl and his colleagues (e.g., Diehl and Kluender, 1989; Diehl *et al.*, 1991; Kingston and Diehl, 1995), suggest that much of the symmetry is a result of talkers producing speech sounds in a manner that exploits general operating characteristics of the auditory system. Direct Realist approaches (e.g., Fowler, 1986, 1996) account for such symmetry in terms of perceptual ''recovery'' of vocal-tract dynamics from the acoustic waveform. Despite this diversity of theoretical accounts, it is clear that effective communication requires perception and production to be in relatively close accord.

One case of symmetry between speech perception and production that has received a considerable amount of empirical attention is the effect of a preceding liquid on stop-consonant perception. Mann (1980) presented listeners with members of a series of synthesized consonant–vowel (CV) syllables varying perceptually from /da/ to /ga/ preceded by natural utterances of either /al/ or /ar/. Subjects identified the CVs as /ga/ more often following /al/ than following /ar/. This shift in responses is complementary to the coarticulatory influences on CV production following /al/ and /ar/. Due to the assimilative nature of coarticulation, the place of vocal-tract occlusion of a stop consonant is more anterior following /al/ than following /ar/. Because the alveolar stop /d/ is produced at an anterior place in the oral cavity, and the velar stop /g/ is produced with a posterior occlusion, /al/ productions result in subsequent CVs being more /da/-like articulatorally and acoustically, while CVs following /ar/ are more /ga/-like. Thus subjects' perceptual responses seem to compensate for acoustic effects of coarticulation; more /ga/ identifications result for CVs following /al/.

Subsequent research has demonstrated this context effect on CV identification for Japanese listeners who are ineffective in distinguishing /al/ and /ar/ (Mann, 1986) and for four-month-old infants (Fowler *et al.*, 1990). These results have been attributed to perceptual representations of, or recovery of, specific vocal-tract dynamics and constraints. According to Mann (1980), ''...speech perception must somehow operate with tacit reference to the dynamics of speech production and its acoustic consequences.''

However, this theoretical view was strongly challenged by results reported by Lotto and Kluender (in press). Participating in a forced-choice identification task like that employed by Mann (1980), subjects identified members of a CV series as /ga/ more often following /al/ than when following /ar/, even when the two syllables were produced by very physically different talkers of opposite gender. Furthermore, the effect remained when the preceding context was a sine-wave caricature modeling a very limited aspect ($F3$ transition) of the /al/ and /ar/ syllables. Lotto and Kluender suggest that ''perceptual compensation for coarticulation'' is probably not due to knowledge of or recovery of specific vocal-tract dynamics because the context effect remains robust for contextual sounds which were clearly not produced by a similar

[a]Correspondence to: Andrew J. Lotto, Department of Psychology, University of Wisconsin, 1202 West Johnson Street, Madison, WI 53706, (608) 262-6110, ajlotto@facstaff.wisc.edu

vocal tract or, indeed, by any vocal tract. They propose that the effect is of general auditory origin and describe the pattern of results as frequency contrast. Redescribed in this light, the results are: following a syllable with a high $F3$-offset frequency (/al/), more low-frequency-$F3$-onset identifications (/ga/) result, and following a syllable with a low $F3$-offset frequency (/ar/), more high-frequency-$F3$-onset identifications (/da/) are obtained. In light of this general framework, Lotto and Kluender (in press) presented subjects with CV syllables similar to those used by Mann (1980) preceded by constant-frequency sine-wave tones with frequencies equivalent to $F3$-offset frequencies of natural /al/ and /ar/ syllables which served as context in the Mann study. The resulting shift in identification functions was actually slightly larger than that obtained by Mann (1980) for natural-speech contexts.

One may assume that accounts which rely on speech-specific mechanisms to accommodate coarticulatory influences would have difficulty explaining these nonspeech results. However, there remain concerns about the validity of research with nonspeech analogs as critical experiments for deciding issues of speech-specificity. Kuhl (1978, 1986a, b) has discussed the possibility that nonspeech stimuli may be accommodated by speech-specific mechanisms with rather broad application. Processes which rely upon abstract kinematic consequences of articulatory dynamics have been proffered, for example, to explain the finding that some subjects are able to hear complexes of sine waves as speech (Remez *et al.*, 1981; Remez *et al.*, 1994).

Even if one does not accept that mechanisms have evolved exclusively for perceiving speech, it is possible that the overlearned nature of speech can affect the perception of quasi-periodic nonspeech sounds. Perceptual heuristics[1] for managing the kinematic characteristics of speech may be developed through the near-constant exposure to these sounds. As a result, contextual sine waves may be processed in a speechlike manner.

In order to test the generality of this context effect and to determine if an account based on general ''frequency contrast'' is viable, an experiment was designed exploiting a nonhuman animal model of speech perception. Previous animal studies of speech perception have been used to assess general auditory processes without confounds of effects of experience and unencumbered by purported speech-specific processes (e.g., Kuhl and Miller, 1975, 1978; Kluender, 1991; Kluender and Lotto, 1994; Dooling *et al.*, 1995). Animals are unlikely recipients of innate speech-specific mechanisms. Consequently, analogous performance on speech tasks for animals and humans, together with the virtue of parsimony, discourages accounts of speech perception which rely on innate representations of gestural dynamics.[2]

I. EXPERIMENT

Four Japanese quail (*Coturnix coturnix japonica*) served as subjects in an experiment designed to test the species-specificity of the contextual effects reported in Mann (1980). Japanese quail have been used successfully in previous experiments concerning the complementary nature of speech perception and production (e.g., Kluender *et al.*, 1987; Kluender, 1991; Kluender and Lotto, 1994). These birds have shown the ability to ''identify'' CV syllables varying in information specifying place of articulation through differential pecking (Kluender *et al.*, 1987). The present experiment tested whether their CV ''labeling'' would be affected by intersyllabic context.

A. Methods

1. Subjects

Four adult Japanese quail served as subjects in the labeling experiment. Free-feed weights ranged from 123 to 154 g.

2. Stimuli

Stimuli were identical to those from the synthesized-speech condition of experiment 2 from Lotto and Kluender (in press). A ten-step series of CV syllables (/da/-/ga/) varying in $F3$-onset frequency was synthesized using the cascade synthesizer described in Klatt (1980). End-point stimuli were based on natural productions of a male talker speaking the syllables in isolation. For these CVs, onset frequency of $F3$ varied from 1800 to 2700 Hz in 100 Hz steps. Then, from onset, $F3$ frequency changed linearly to a steady-state value of 2450 Hz over an 80-ms transition.[3] Amplitude of $F3$ was approximately 3 dB less intense than $F1$ and $F2$ at onset. All other synthesizer parameters remained constant across members of the series. Frequency of the first formant ($F1$) increased from 300 to 750 Hz and second-formant ($F2$) frequency decreased from 1650 to 1200 Hz over 80 ms. Fundamental frequency ($f0$) was 110 Hz from onset until decreasing linearly to 95 Hz over the last 50 ms. Total stimulus duration of synthesized CVs was 250 ms.

Three syllables serving as preceding context were also synthesized with values based on utterances of the same talker upon whose productions the CV series was modeled. All three syllables were 250 ms in duration and had a constant $f0$ of 110 Hz. Two of the preceding-context syllables were synthesized versions of /al/ and /ar/. For each syllable, the first 100 ms consisted of the same steady-state vowel. The frequencies of the first four formants during this steady state were 750, 1200, 2450, and 2850 Hz, respectively. Following this 100-ms vowel were 150-ms linear formant transitions. Offset frequencies for the first four formants in the /al/ syllable were 564, 956, 2700, and 2850 Hz, respectively. For the /ar/ syllable these values were 549, 1517, 1600, and 2850 Hz. The third preceding context was a 250-ms steady-state vowel /a/ synthesized with the same parameters as for the vowel in each VC.

Stimuli were synthesized with 12-bit resolution at a 10-kHz sampling rate, matched in rms level and stored on computer disk. Stimulus presentation was under control of an 80386 computer. Following D/A conversion (Ariel DSP-16), stimuli were low-pass filtered (4.8-kHz cutoff frequency, Frequency Devices #677), amplified, and presented to subjects via a single 13-cm speaker (Peerless 1592) in a tuned enclosure providing flat frequency response from 40 to 5000 Hz. Sound level was calibrated by placing a small sound-level meter (Bruel & Kjaer 2232) in the chamber with the

microphone positioned at approximately the same height and distance from the speaker as the performing bird's head.

3. Procedure

Quail first were trained by means of operant procedures to discriminate stimuli from each end of the CV series. For two birds (quail 1 and quail 2), CVs with low $F3$-onset frequencies (1800 and 1900 Hz) signaled positive reinforcement (/ga/-positive), while for the other two birds (quail 3 and quail 4), CVs with high $F3$-onset frequencies (2600 and 2700 Hz) were reinforced (/da/-positive). Following 18 to 22 h of food deprivation (adjusted to each bird individually for optimal performance[4]), birds were placed in a soundproof operant chamber (Industrial Acoustics Corp., model AC1) inside a larger single-wall soundproof booth (Suttle Acoustics Corp.). In a go/no-go identification task, birds pecked a single lighted 1.2-cm-square key located 15 cm above the floor and centered below the speaker. Stimuli were presented, responses were recorded, and reinforcement was controlled by an 80386 microcomputer.

For three quail, the training sequence was identical. CVs were presented in the three contexts (/a/, /al/, and /ar/) from the beginning of training. For the other quail (quail 4), CVs were presented in isolation until the bird's peck ratios reached a performance criterion (10:1 peck ratio for positive versus negative stimuli), then training was continued with inclusion of contextual sounds. This alternative procedure for quail 4 resulted in no discernible difference in the final data.

During training with contextual sounds, stimuli consisted of a disyllable including one of the three contexts (/a/, /al/, or /ar/) followed by a 50-ms silent interval (typical of natural productions) and then one of the four training CV syllables (with $F3$ frequencies of 1800, 1900, 2600, and 2700 Hz). Appending of syllables was accomplished digitally online during the experiment. On each trial, a disyllable was repeatedly presented once per 1550 ms at an average peak level of 70 dB SPL. (During the initial training of quail 4, single CVs were repeated once per 1550 ms.) On a trial-by-trial basis, the intensity of the disyllable (or single CV) was varied randomly from 70 dB by ± 0–5 dB [mean=70 dB SPL] through a computer-controlled digital attenuator (Analog Devices 7111). Average duration of each trial was 30 s, varying geometrically from 10 to 65 s. Intertrial interval was 15 s. Responses to positive stimuli were reinforced on a variable-interval schedule by 1.5–2.5 s access to food from a hopper beneath the peck key. Duration of reinforcement was also adjusted for each bird for consistent performance. Average interval to reinforcement was 30 s (10–65 s), so that positive stimuli were reinforced on an average of once per trial. Note that when a trial was long (e.g., 57 or 65 s) and times to reinforcement were short (e.g., 10 or 12 s), reinforcement was available more than once. Likewise, on shorter positive trials reinforcement did not become available if time to reinforcement was longer than the trial. Any reinforcement interval that did not expire during one positive trial carried over to the next positive trial. Such intermittent reinforcement encouraged consistent peck rates during later nonreinforced testing trials. During negative trials, birds were required to refrain from pecking for 5 s for presentation of the stimulus, and the trial to be terminated. This procedure has been used successfully to train Japanese quail in similar speech perception tasks (Kluender, 1991; Kluender and Lotto, 1994).

Following magazine training and autoshaping procedures, reinforcement contingencies were gradually introduced over a 1-week period in sessions of 60 to 72 trials each. During that first week the average amplitude of the stimuli was increased from 50 to 70 dB SPL in order to introduce sound without startling the birds. (Following the attainment of a performance criterion, contextual sounds were gradually presented to quail 4. The average amplitude of these context sounds was increased from about 40 to 70 dB over a 2-week period.) Also during this first week: average trial duration increased from 5 to 30 s; intertrial interval decreased from 40 to 15 s; average time to reinforcement was increased from 5 to 30 s; access to the food hopper was decreased from 4.0 to 2.0 s; and ratio of positive to negative trials decreased from 4:1 to 1:1.

All birds learned quickly to respond differentially to high $F3$-onset frequency versus low $F3$-onset frequency CVs, pecking at least twice as often to positive stimuli versus negative stimuli by the end of 50 days of training (3600 trials). One bird (quail 2) was pecking at a 2:1 ratio after only 23 days of training (1656 trials). Whereas Kluender and Lotto (1994) found 2:1 performance after only 20 days of training for a voiced/voiceless distinction, the task for the current birds may be considered more difficult in that reinforced discrimination must be made solely on a small change in relative amplitudes of harmonics as opposed to the multifarious cues which result from changes in voice-onset time. Birds continued to train with the four extreme CVs in three contexts until they achieved 10:1 performance for positive versus negative stimuli. Attaining this level of performance required between 140 and 163 days of training (10 080 to 11 160 trials). (In addition, quail 4 needed 88 days (6336 trials) to reach 10:1 performance after contextual sounds were added.)

Following training, the four birds were tested on novel CV syllables with intermediate $F3$-onset frequencies ranging from 2000 to 2500 Hz (100-Hz steps) in all three contexts (/a/, /al/, and /ar/). Due to uncontrollable scheduling conflicts, the number of testing trials with each novel disyllable varied between birds. Quail 1 was presented all six intermediate CVs with each context eight times across 16 days of testing. Quail 2 responded to each disyllable 10 times over 20 days; nine times over 18 days for quail 3; 20 times over 40 days for quail 4. During a single test session, nine of the possible 18 novel disyllables (six CVs$\times$three contexts) were presented, each on one 30-s trial. During the presentation of novel disyllables, no contingencies were in effect. Birds neither received food reinforcement nor needed to refrain from pecking for presentation to terminate after 30 s. Each testing session of 69 total trials began with 15 reinforced trials with training disyllables followed by nine nonreinforced trials with novel disyllables interspersed amongst 45 reinforced

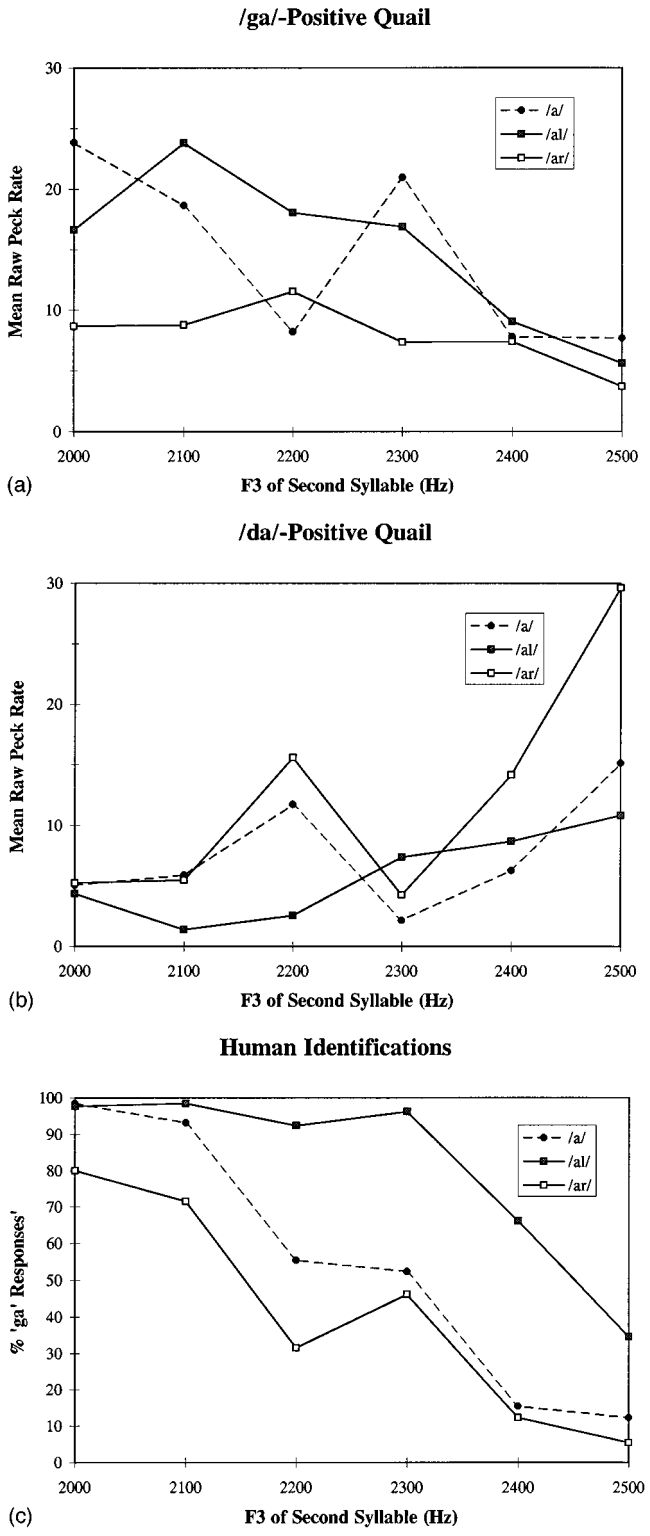


FIG. 1. Mean raw peck rates for each of the six novel /Ca/ stimuli in each of the three context conditions. (a) Average peck rates for birds trained to peck to /ga/ stimuli. (b) Average peck rates for birds trained to peck to /da/ stimuli. (c) Mean percentage of /ga/ identifications for human listeners from Lotto and Kluender (in press).

trials with training stimuli. Novel trials could not occur until 15 training stimuli trials had been presented to assure that each bird ''settled into'' the task before responding to test stimuli. For all training and testing sessions, stimuli were randomly ordered for each bird.

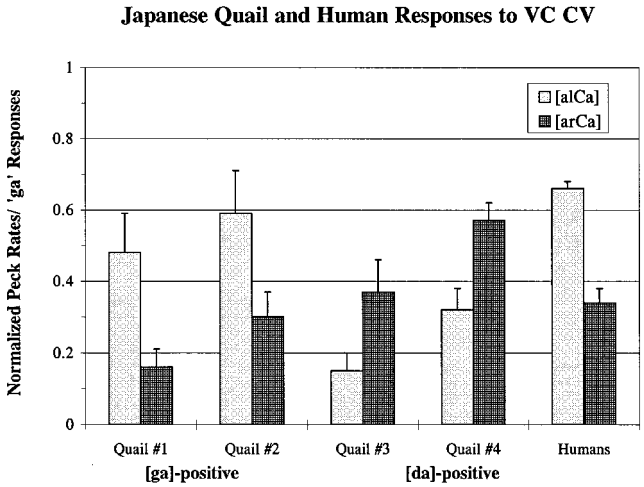


FIG. 2. Histogram contains normalized peck rates (percentages of range of mean peck rates) along with attendant standard errors for intermediate CV syllables in the context of /al/ and /ar/. Human data is from Lotto and Kluender (in press) synthesized-speech condition of experiment 2. Represented here are the percent /ga/ responses to those CVs which served as test stimuli in the current quail experiment. Mean percentages were normalized so that they summed to 1.00 to put them on a scale similar to that used for peck rates.

B. Results and discussion

For each bird, raw pecks were collected for each test trial. These test trials were all fixed at a duration of 30 ms. Figure 1 displays the raw peck data for each of the six novel stimuli in each of the three context conditions. Figure 1(a) is a plot of the data averaged across the responses of the two birds trained to peck to /ga/ (low-frequency $F3$) stimuli and Fig. 1(b) is a plot of the data collected from the birds trained to peck to /da/ (high-frequency $F3$) stimuli. Figure 1(c) is a plot of human identification functions from Lotto and Kluender (in press) for the novel stimuli. Note that there is a nonlinearity for the birds and humans for the CV with an $F3$ onset of 2200 Hz. This stimulus differs only in $F3$-onset frequency from the other stimuli and experiments are being conducted to discover the reason for this nonlinearity. It is interesting that there is such qualitative agreement in the functions for two different species, presumably performing different tasks.

An alternative data representation was also calculated to deal with the inherent variance arising from the different peck rates of individual birds. Total pecks to each disyllable were summated for each ''run'' through all of the novel stimuli, i.e., the data for all 18 novel disyllables collected across two days. Mean peck rates (pecks per 30-s trial) were calculated for each of the three contexts (/a/, /al/, /ar/). These means were then transformed into percentages of the mean range. The lowest mean for each bird was subtracted from each mean and the result was divided by the range of mean scores (maximum mean−minimum mean). For example, for quail 3, the highest mean peck rate for any context was 39.33 pecks/30 s. This was the mean peck rate across all novel stimuli with the /ar/ context on its first testing run. The minimum mean peck rate for this bird was 0.0 pecks/30 s for all novel stimuli in the /al/ context on the sixth run. The nor-

TABLE I. Results of *t* tests and Wilcoxon signed-rank tests on normalized peck rates for each subject.

Bird	/al/ mean	/ar/ mean	*t* value	*p* value	Wilcoxon	*p* value
Quail 1	0.48	0.16	3.57	0.009	0.0	0.008
Quail 2	0.59	0.30	2.50	0.034	8.0	0.047
Quail 3	0.15	0.37	2.74	0.025	4.0	0.028
Quail 4	0.32	0.57	6.41	0.000	6.0	0.000

malized mean peck rate for the /ar/ context of the third run would be

$$\frac{(\text{mean for /ar/ of third run})-(\text{minimum mean})}{(\text{maximum mean})-(\text{minimum mean})}$$

$$=\frac{(20.83)-(0.00)}{(39.33)-(0.00)}=0.53.$$

This transform normalizes the mean peck rates in terms of the range of peck rates for an individual bird, thus minimizing variance that arises from the fact that some birds are ''heavier'' peckers than others. It has been determined through Monte Carlo simulations that such a range transform increases power without increasing the likelihood of type I error (Bush *et al.*, 1993). Normalized mean peck rates with attendant standard errors for the /al/ and /ar/ contexts are presented in Fig. 2. Quails 1 and 2 were trained to peck to /ga/ stimuli and show increased pecking in the /al/ condition. Conversely, quails 3 and 4, which were trained to peck to the /da/ end points, pecked more to novel stimuli following /ar/. For comparison, data from adult humans hearing the same stimuli (Lotto and Kluender, in press) are presented. These data were normalized by dividing the number of /ga/ responses in each condition by the total number of /ga/ responses across conditions.

Matched-pairs *t* tests were computed for the difference between normalized peck rates in /al/ and /ar/ contexts for each bird. Runs through the stimuli (i.e., 2-day sessions in which each stimulus is heard once) were treated as the random variable for the tests. For each bird, normalized peck rates were significantly different in /al/ than in /ar/ contexts.[5] Table I presents the outcomes of each *t* test along with the outcomes of independently conducted nonparametric Wilcoxon signed-rank tests. For birds trained to peck in response to CVs with high $F3$-onset frequencies (/da/-positive: quail 3 and quail 4), peck rates increased for CV syllables following /ar/. For birds trained to peck to CVs with low $F3$-onset frequencies (/ga/-positive: quail 1 and quail 2), peck rates increased for CV syllables following /al/. This pattern of responses is analogous to that of human listeners in identification tasks, who respond with /ga/ identifications more often following /al/ than following /ar/ (Mann, 1980, 1986; Fowler *et al.*, 1990; Lotto and Kluender, in press).

II. GENERAL DISCUSSION

The experiment in this report was designed to test the generality of the effect of preceding liquid identity on stop-consonant identification first reported in Mann (1980). The fact that nonhuman animals demonstrated effects of context on labeling of CV syllables similar to that found for human listeners suggests that this effect is of general auditory origin. Japanese quail are unlikely to possess processes designed specifically for the domain of human speech. These birds also had no chance to learn the covariation of formant frequencies caused by assimilative processes of coarticulation. As a matter of parsimony, one is thrust into the theoretical position to recommend that ''perceptual compensation for coarticulation'' is the outcome of a rather general auditory process.

Consequences of this general process can be described as contrastive. Quail trained to peck to CVs with *low* $F3$-onset frequencies (/ga/-positive) pecked more to intermediate values of $F3$-onset frequency when CVs were preceded by a syllable with *a high*-frequency $F3$ offset (/al/). Quails trained to peck to CVs with *high* $F3$-onset frequencies (/da/-positive) pecked more to intermediate values of $F3$-onset frequency when CVs were preceded by a syllable with a low-frequency $F3$ offset (/ar/). This contrastive pattern has now been shown for adult human English speakers (Mann, 1980; Lotto and Kluender, in press), adult human Japanese speakers (Mann, 1986), 4-month-old infants (Fowler *et al.*, 1990), and Japanese quail.

In fact, contrastive perceptual effects, as those noted here for human and bird listeners, are epidemic. In the visual system, perceptual contrast has been described for lightness perception (Koffka, 1935; Wallach, 1948), line orientation (Gibson, 1933, 1937; Gibson and Radner, 1937), size, position, and curvature (Kohler and Wallach, 1944), spatial frequency (Blakemore and Sutton, 1969), depth (Ames, 1935; Kohler and Emery, 1947; Bergman and Gibson, 1959), and color (Cathcart and Dawson, 1928–1929). Contrastive effects have been witnessed in tempo of behavior (Cathcart and Dawson, 1927–1928) and lifting of weights (Guilford and Park, 1931; Sherif *et al.*, 1958). In audition, frequency contrast has been demonstrated (Cathcart and Dawson, 1928–1929; Christman, 1954) as has contrast in lateralization of a sound (Flügel, 1920–1921).

In speech perception literature, subjects' responses often can be described in terms of contrast of some type. For example, identification boundaries for members of a stop/glide series varying in transition duration shift toward longer transitions (i.e., more *short*-transition responses) when syllable duration is *increased* (Miller and Liberman, 1979). Similarly, syllable-final consonants are judged more often to be voiced (*shorter* silent interval) when the duration of the preceding vowel is *increased* (Denes, 1955; Raphael, 1972; Port and Dalby, 1982). In each of these cases, including VC CV context effects described in this report, contrastive perception appears to compensate for articulatory regularities.

Given the ubiquity of contrastive perception, one may conjecture that it serves an important adaptive purpose. But what purpose is served by a general perceptual characteristic which causes, at first blush, seemingly nonveridical percepts? Why would a Japanese quail benefit from a process which alters perceived frequency of spectral components depending on context?[6] Part of the answer may lie in the remarkable symmetry between speech perception and production that was noted in the Introduction.

Due to the variables of inertia and mass, physical systems tend to be assimilative. The configuration of a system at time t is significantly constrained by its configuration at time $t-1$. The set of possible transformations from time t to $t-1$ is also limited (e.g., by a constraint of rigidity; Ullman, 1984). Perceptual systems have developed in an environment governed by particular physical laws and it is probable that perceptual processes respect these laws. This is a form of an argument advanced by Shepard (1984) under the name ''Psychophysical Complementarity.''

Because very rapid change is the exception for physical systems with mass and inertia, signs of change are emphasized through the processes of perceptual contrast. This is due, perhaps, to the ecological importance of rapid change, especially as it relates to the default of continuance in physical systems. For example, lightness contrast emphasizes boundaries at which there is a rapid change in luminance. This may help establish the borders of separate objects. This is similar to accounts of perceptual contrast which assume that perception is referenced to or ''anchored'' upon some previously presented standard (e.g., Helson, 1964; Warren, 1985).

As for speech communication, coarticulation is due, at least in part, to physical constraints on articulators and to dynamic variables of mass and inertia (Ostry *et al.*, 1996). Vocal-tract shape changes relatively smoothly over time. At conversational speaking rates, articulators often undershoot target articulations which are produced in clear speech (Lindblom, 1963). Because it is a physical system, signals generated by a vocal tract are perceived in a contrastive manner by both humans and Japanese quail. The resulting symmetry of production and perception is not serendipitous, but is a consequence of organisms having evolved to interact with physical systems which are constrained across time. In this light, the results with Japanese quail are not too surprising. The physical environment of the antecedents of the quail likely resemble that of early hominids and analogous perceptual solutions probably have developed.

This account is, for now, too superficial to qualify as a full theoretical account. However, it does have some concepts in common with major theories in speech communication. Along with Motor Theory, it acknowledges that articulators are highly constrained across time and that the resultant signal is largely a product of articulatory constraints and dynamics. It mirrors Direct Realism in its appreciation for the ecology of sound and the sources which produce it. It shares with Auditorist theories a tenet that general auditory processes are culpable for many of the phenomena of speech perception.

The disjunctions between this account and those which precede it must also be given consideration. It denies the necessity of tacit gestural representations and speech-specific processes as proposed by Motor Theory. It cannot be characterized as Direct Realism in as much as the proposed percepts are not veridical and need not correspond to real ''objects.'' And finally, its emphasis is shifted from traditional Auditorist accounts. In the present account, the symmetry between speech perception and production arises from the perceptual system accommodating the constraints on physical systems such as articulators, and not from the dictatorial demands of the operating characteristics of the auditory system.

As facile as this account may be, it addresses the mounting evidence of similar perceptual behavior for nonhuman animals and humans with speech stimuli (e.g., Kluender *et al.*, 1987; Kluender, 1991; Kluender and Lotto, 1994; Dooling *et al.*, 1995). In the present case, Japanese quail have shown context effects for speech sounds; an effect previously described as ''perceptual compensation for coarticulation'' (e.g., Mann 1980). As impressive as the symmetry between speech perception and production is, it is not an exclusively human achievement.

ACKNOWLEDGMENTS

This work was supported by NIDCD Grant No. DC-00719 and NSF Young Investigator Award DBS-9258482 to the second author and by Sigma Xi Dissertation Research Award to the first author. Some of the data were presented at the International Congress of Phonetic Sciences, August 1995, in Stockholm.

[1]''Heuristic'' is intended only to refer to a class of algorithms which offer potentially fallible solutions to problems, but are nevertheless useful in most situations. The term is often expanded with theoretical content in cognitive science.

[2]Comparable perceptual behavior of humans and animals should not bother purveyors of Direct Realist theories. To the contrary, since, by this view, the information specifying articulatory dynamics is inherent in acoustic signals produced by vocal tracts, nonhuman animals should be able to recover this information (see, e.g., Fowler, 1996). However, the results of Lotto and Kluender (in press) should be particularly troubling as they demonstrate contextual effects for sounds which clearly originate from different sources.

[3]Formant transitions of 80-ms duration may seem to be rather long, but these were measured from natural productions and are comparable to the 100-ms transitions used by Mann (1980).

[4]Optimal performance was defined as the highest ratio of pecks to positive versus negative stimuli. Birds were idiosyncratic with regard to the amount of deprivation that resulted in the most stable performance, and weights ranged from 80% to 95% of free-feed weight at the time of training/testing.

[5]For all birds, mean peck rates to CVs in the /a/ context fell between those for CVs in the /al/ and /ar/ contexts.

[6]It should be noted that changes in formant frequencies are not changes in frequency, *per se*. Harmonics remain at the same frequency, but there is a change in the relative energy across the spectrum. It may be more correct to state that the perception of relative energy is context-dependent.

Ames, A. (**1935**). ''Aneiseikonia—a factor in the functioning of vision,'' Am. J. Opthamology **28**, 248–262.

Bergman, R., and Gibson, J. J. (**1959**). ''The negative aftereffect of a surface slanted in the third dimension,'' Am. J. Psychol. **72,** 364–374.

Blakemore, C., and Sutton, P. (**1969**). ''Size adaptation: a new aftereffect,'' Science **166,** 245–247.

Bush, L. K., Hess, U., and Wolford, G. (**1993**). ''Transformations for within-subject designs: A Monte Carlo investigation,'' Psychol. Bull. **113,** 566–579.

Cathcart, E. P., and Dawson, S. (**1927–1928**). ''Persistence: A characteristic of remembering,'' Brit. J. Psychol. **18**, 262–275.

Cathcart, E. P., and Dawson, S. (**1928–1929**). ''Persistence (2),'' Brit. J. Psychol. **19**, 343–356.

Christman, R. J. (**1954**). ''Shifts in pitch as a function of prolonged stimulation with pure tones,'' Am. J. Psychol. **67,** 484–491.

Diehl, R. L., and Kluender, K. R. (**1989**). ''On the objects of speech perception,'' Ecol. Psychol. **1**, 121–144.

Diehl, R. L., Kluender, K. R., Walsh, M. A., and Parker, E. M. (**1991**). ''Auditory enhancement in speech perception and phonology,'' in *Cogni-*

tion and the Symbolic Processes: Applied and Ecological Perspectives, edited by R. Hoffman and D. Palermo (Erlbaum, Hillsdale, NJ), pp. 59–75.

Denes, P. (**1955**). ''Effect of duration on the perception of voicing,'' J. Acoust. Soc. Am. **27,** 761–764.

Dooling, R. J., Best, C. T., and Brown, S. D. (**1995**). ''Discrimination of synthetic full-formant and sinewave /ra-la/ continua by budgerigars (*Melopsittacus undulatus*) and zebra finches (*Taeniopygia guttata*),'' J. Acoust. Soc. Am. **97,** 1839–1846.

Flügel, J. C. (**1920–1921**). ''On local fatigue in the auditory system,'' Brit. J. Psychol. **11**, 105–134.

Fowler, C. A. (**1986**). ''An event approach to the study of speech perception from a direct-realist perspective,'' J. Phon. **14**, 3–28.

Fowler, C. A. (**1996**). ''Listeners do hear sounds, not tongues,'' J. Acoust. Soc. Am. **99,** 1730–1741.

Fowler, C. A., Best, C. T., and McRoberts, G. W. (**1990**). ''Young infants' perception of liquid coarticulatory influences on following stop consonants,'' Percept. Psychophys. **48,** 559–570.

Gibson, J. J. (**1933**). ''Adaptation, after-effect and contrast in the perception of curved lines,'' J. Exp. Psychol. **16,** 1–31.

Gibson, J. J. (**1937**). ''Adaptation with negative after-effect,'' Psychol. Rev. **44,** 222–244.

Gibson, J. J., and Radner, M. (**1937**). ''Adaptation, after-effect, and contrast in the perception of tilted lines. I. Quantitative Studies,'' J. Exp. Psychol. **20,** 453–467.

Guilford, J. P., and Park, D. G. (**1931**). ''The effect of interpolated weights upon comparative judgments,'' Am. J. Psychol. **43,** 589–599.

Helson, H. (**1964**). *Adaptation-Level Theory: An Experimental and Systematic Approach to Behavior* (Harper & Row, New York).

Kingston, J., and Diehl, R. L. (**1995**). ''Intermediate properties in the perception of distinctive feature values,'' in *Papers in Laboratory Phonology IV*, edited by A. Arvaniti and B. Connell (Cambridge U.P., Cambridge, England).

Klatt, D. H. (**1980**). ''Software for a cascade/parallel formant synthesizer,'' J. Acoust. Soc. Am. **67,** 971–995.

Kluender, K. R. (**1991**). ''Effects of first formant onset properties on voicing judgments result from processes not specific to humans,'' J. Acoust. Soc. Am. **90,** 83–96.

Kluender, K. R., Diehl, R. L., and Killeen, P. (**1987**). ''Japanese quail can learn phonetic categories,'' Science **237,** 1195–1197.

Kluender, K. R., and Lotto, A. J. (**1994**). ''Effects of first formant onset frequency on [-voice] judgments result from general auditory processes not specific to humans,'' J. Acoust. Soc. Am. **95,** 1044–1052.

Koffka, K. (**1935**). *Principles of Gestalt Psychology* (Harcourt, Brace & World, New York).

Kohler, W., and Emery, D. A. (**1947**). ''Figural aftereffects in the third dimension of visual space,'' Am. J. Psychol. **60,** 159–201.

Kohler, W., and Wallach, H. (**1944**). ''Figural aftereffects: An investigation of visual processes,'' Proc. Am. Philos. Soc. **88**, 269–357.

Kuhl, P. K. (**1978**). ''Predispositions for the perception of speech-sound categories: A species-specific phenomenon,'' in *Communicative and Cognitive Abilities-Early Behavioral Assessment*, edited by F. D. Minifie and L. L. Lloyd (University Park Press, Baltimore), pp. 229–255.

Kuhl, P. K. (**1986a**). ''Theoretical contributions of tests on animals to the special-mechanisms debate in speech,'' Exp. Biol. **45,** 233–265.

Kuhl, P. K. (**1986b**). ''The special-mechanisms debate in speech research: Categorization tests on animals and infants,'' in *Categorical Perception*, edited by S. Harnad (Cambridge U.P., Cambridge, MA), pp. 355–386.

Kuhl, P. K., and Miller, J. D. (**1975**). ''Speech perception by the chinchilla: Voiced-voiceless distinction in alveolar plosive consonants,'' Science **190,** 69–72.

Kuhl, P. K., and Miller, J. D. (**1978**). ''Speech perception by the chinchilla: Identification functions for synthetic VOT stimuli,'' J. Acoust. Soc. Am. **63,** 905–917.

Liberman, A. M., and Mattingly, I. G. (**1985**). ''The motor theory of speech perception revised,'' Cognition **21**, 1–36.

Lindblom, B. (**1963**). ''Spectrographic study of vowel reduction,'' J. Acoust. Soc. Am. **35,** 1773–1781.

Lotto, A. J., and Kluender, K. R. (**in press**). ''General contrast effects in speech perception: Effect of preceding liquid on stop consonant identification,'' Percept. Psychophys.

Mann, V. A. (**1980**). ''Influence of preceding liquid on stop-consonant perception,'' Percept. Psychophys. **28,** 407–412.

Mann, V. A. (**1986**). ''Distinguishing universal and language-dependent levels of speech perception: Evidence from Japanese listeners perception of English 'l' and 'r','' Cognition **24**, 169–196.

Miller, J. L., and Liberman, A. M. (**1979**). ''Some effects of later-occurring information on the perception of stop consonant and semivowel,'' Percept. Psychophys. **25,** 457–465.

Ostry, D. J., Gribble, P. L., and Gracco, V. L. (**1996**). ''Coarticulation of jaw movements in speech production: Is context sensitivity in speech kinematics centrally planned?,'' J. Neurosci. **16,** 1570–1579.

Port, R. F., and Dalby, J. (**1982**). ''Consonant/vowel ratio as a cue for voicing in English,'' Percept. Psychophys. **32,** 141–152.

Raphael, L. F. (**1972**). ''Preceding vowel duration as a cue to the perception of the voicing characteristics of word-final consonants in English,'' J. Acoust. Soc. Am. **51,** 1296–1303.

Remez, R. E., Rubin, P. E., Berns, S. M., Pardo, J. S., and Lang, J. M. (**1994**). ''On the perceptual organization of speech,'' Psychol. Rev. **101,** 129–156.

Remez, R. E., Rubin, P. E., Carrell, D. B., and Carrell, T. D. (**1981**). ''Speech perception without traditional speech cues,'' Science **212,** 947–950.

Shepard, R. N. (**1984**). ''Ecological constraints on internal representation: Resonant kinematics of perceiving, imagining, thinking, and dreaming,'' Psychol. Rev. **91,** 417–447.

Sherif, M., Taub, D., and Hovland, C. I. (**1958**). ''Assimilation and contrast effects of anchoring stimuli on judgments,'' J. Exp. Psychol. **55,** 150–155.

Ullman, S. (**1984**). ''Maximizing rigidity: the incremental recovery of 3-D structure from rigid and nonrigid motion,'' Perception **13**, 255–274.

Wallach, H. (**1948**). ''Brightness constancy and the nature of achromatic colors,'' J. Exp. Psychol. **38,** 310–324.

Warren, R. M. (**1985**). ''Criterion shift rule and perceptual homeostasis,'' Psychol. Rev. **92,** 574–584.

Speech enhancement using state-based estimation and sinusoidal modeling

Michael E. Deisher[a] and Andreas S. Spanias[b]
Department of Electrical Engineering, Telecommunications Research Center, Arizona State University, Tempe, Arizona 85287-7206

(Received 23 January 1996; revised 4 January 1997; accepted for publication 14 March 1997)

A procedure for estimating the parameters of a sinusoidal model from speech corrupted by additive noise is described. An *approximate harmonic* representation is used wherein voiced speech is represented by a set of sine waves at multiples of the fundamental frequency and several additional components at frequencies near each harmonic. Amplitudes and phases of the sinusoidal components are estimated using a state-based technique that employs hidden Markov models (HMMs) to classify speech and noise spectra. Voicing and fundamental frequency are determined using an analysis-by-synthesis approach. Simulation results are presented, comparing the performance of the proposed algorithm to that of the standard HMM-based minimum mean square error (MMSE) estimator. The proposed method was found to reduce the structured residual noise associated with HMM-based algorithms. © *1997 Acoustical Society of America.*
[S0001-4966(97)02407-7]

PACS numbers: 43.72.Ew [JS]

INTRODUCTION

Speech communication in mobile environments is often difficult due to high-energy ambient noise. The reduction in speech *quality* due to noise is known to cause listener fatigue. Moreover, speech *intelligibility* is severely reduced when low-energy, perceptually important speech is masked by high-energy noise. Speech enhancement algorithms attempt to improve these perceptual aspects of degraded speech.

In addition to improving speech quality or intelligibility for the human listener, speech enhancement preprocessors can improve the performance of other speech processing algorithms. For example, the accuracy of speech recognition algorithms used for ''hands-free'' dialing of mobile cellular telephones is severely reduced when speech is corrupted by background noise. In this situation, a speech enhancement preprocessor can be added to improve recognition accuracy, In addition, speech compression algorithms used in digital cellular telephones perform poorly in noisy environments, especially when coding at low bit rates. A speech enhancement preprocessor can be employed in this case to decrease loss in the coder.

The sinusoidal model represents speech over short time intervals by a finite set of sinusoids that capture most of the signal energy. This model has been used in speech coding applications because it provides an accurate, compact representation for speech.[1] Sinusoidal analysis/synthesis has also been successfully applied to co-channel talker interference suppression.[2] This model is also of interest for the enhancement of voiced speech corrupted by additive noise. In cases where the interference has a fairly flat spectrum, the model tends to select the components with the highest SNR (signal-to-noise ratio). These components can be estimated more accurately than those ''buried'' in noise. In addition, the model naturally rejects interference whose energy is concentrated between the frequencies of the selected components.

In this paper, we describe a procedure for estimating sinusoidal parameters (amplitudes, phases, and frequencies) from speech corrupted by additive noise. The technique used to obtain amplitude and phase estimates is the hidden Markov model (HMM) based minimum mean square error (MMSE) estimator.[3] The HMM approach is attractive here because (a) it is well suited to nonstationary noise, (b) it requires only one microphone, and (c) it may be trained for a variety of noise conditions. Unlike the original HMM-based estimator,[3] however, the algorithm described in this paper estimates a *reduced* set of frequency-domain parameters. Interestingly, the MMSE bit allocation scheme used in adaptive transform coding (ATC)[4] results in many nonharmonic components of voiced speech being set to zero (i.e., represented by zero bits). However, the algorithm presented here differs in that a hidden Markov model is used to describe the noise process, interharmonic components are *always* set to zero, and surviving spectral components are, of course, not quantized. The Soft-Decision Sine Wave Vector Quantizer (SDSWVQ) proposed by Quatieri and McAulay[5] makes use of the sinusoidal model for noise reduction. However, the SDSWVQ was designed specifically for white, stationary noise. Furthermore, it is dependent on accurate estimation of the fundamental frequency during voiced speech segments. The method proposed here uses hidden Markov modeling to cope with variations in noise between speech pauses and utilizes an *approximate* harmonic representation to reduce dependence on accurate fundamental frequency information.

The rest of the paper is organized as follows. The sinusoidal model and its application to noisy speech are described in Sec. II. Section III contains a short summary of the HMM-based speech enhancement approach. The proposed

[a]Electronic mail: Michael_Diesher@ccm.jf.intel.com
[b]Electronic mail: spanias@asu.edu

speech enhancement system is described in Sec. IV. Section V presents the experimental results and is followed by concluding remarks in Sec. VI.

I. THE SINUSOIDAL MODEL

The sinusoidal speech model[1] represents short speech segments as a sum of weighted sinusoids. Specifically, the ith sample of speech within a 20- to 30-ms block is represented by

$$y(n)=\sum_{\kappa=1}^{K_s} A_\kappa \cos(\omega_\kappa n+\phi_\kappa), \tag{1}$$

where K_s is the number of sinusoids used in the representation and A_κ and ϕ_κ are the amplitude and phase associated with the κth frequency track ω_κ. The parameters $\{A_\kappa,\omega_\kappa,\phi_\kappa\}$ provide a representation for one block of speech. The amplitudes $\{A_\kappa\}$ and phases $\{\phi_\kappa\}$ are typically measured from a high-resolution discrete Fourier transform (DFT). Although the frequencies ω_κ need not be harmonically related, the model is often simplified by assuming a harmonic relationship during voiced speech and an equidistant relationship during unvoiced speech, especially in speech coding applications where a compact representation is desired.

When the harmonic sinusoidal model is applied in the presence of additive noise [with samples denoted $v(n)$] the parameters $\{A_\kappa,\omega_0,\phi_\kappa,s\}$ must be estimated from noisy observations, $z(n)=y(n)+v(n)$. The variable s represents the ternary voicing state (voiced, unvoiced, or nonspeech). Estimation of the fundamental frequency from noisy speech is a difficult problem. However, the synthesis model in a speech enhancement application need not be as compact as the one used in sinusoidal transform coding (STC). In fact, a rich parametric set *must* be chosen to ensure good quality speech and avoid ''vocoding'' noise. The approach taken here to weaken the dependence on an accurate fundamental frequency measurement and suppress vocoding noise is to choose several additional sinusoidal components surrounding the harmonics, namely all DFT components in the intervals

$$\{[\omega_0(\kappa-\rho/2),\omega_0(\kappa+\rho/2)]\}_\kappa, \tag{2}$$

where $0\leqslant\rho\leqslant 1$. During unvoiced speech or pauses, ρ is set to one. This corresponds to using all DFT components. During voiced speech, the value of ρ must be chosen to trade off the effects of fundamental frequency and model errors versus the effects of interharmonic noise. Selection of ρ is discussed further in Sec. IV.

In addition to the fundamental frequency, amplitudes and phases must also be determined from noisy measurements. Let $V_t(k)$, $Y_t(k)$, and $Z_t(k)$ be the kth DFT components of the noise, speech, and noisy speech, respectively. The subscript $(\cdot)_t$ is an integer block index. One possible estimate of $Y_t(k)=\frac{1}{2}A_{t,\kappa}\angle\,\phi_{t,\kappa}$ is the minimum mean square error (MMSE) estimate,

$$\widehat{Y_t(k)}=E\{Y_t(k)|Z_t(k)\}. \tag{3}$$

To calculate this estimate, the following assumptions are made. Clean speech and noise blocks ($\mathbf{y_t}$ and $\mathbf{v_t}$) are assumed to be realizations of zero mean, stationary random processes. The DFT components of speech $\{Y_t(k)\}_{k=0}^{K-1}$ and noise $\{V_t(k)\}_{k=0}^{K-1}$ are each assumed to have jointly Gaussian probability densities. Since DFT components are uncorrelated, the assumption that they are jointly Gaussian is equivalent to assuming that they are independent (wrt frequency index k). Finally, the clean speech and noise are assumed to be statistically independent of one another. Under these assumptions, the MMSE estimate stated in (3) is

$$\widehat{Y_t(k)}=H_t(k)Z_t(k)=\frac{E\{|Y_t(k)|^2\}}{E\{|Y_t(k)|^2\}+E\{|V_t(k)|^2\}}\,Z_t(k). \tag{4}$$

Here, H_t is known as an ''approximate Wiener filter.''[6] Equation (4) is not directly applicable because the second-order statistics of a particular block of speech and noise are not usually known in advance. Fortunately, the HMM-based speech enhancement approach provides a way to circumvent this problem.

II. HMM-BASED ENHANCEMENT APPROACH

Hidden Markov models have proven to be quite useful in modeling the probability density functions of speech parameters in voice recognition. The HMMs are also useful models for other time-varying signals including many classes of noise sources. The HMM-based speech enhancement approach assumes that both speech and noise are well modeled by HMMs. This method requires training two Gaussian, autoregressive, *ergodic* HMMs (i.e., HMMs where all state transitions are allowed): one for clean speech and another for noise. Typically, a large amount of training data is required to sufficiently train the models. Given the trained models, a MMSE estimation strategy is then pursued.[3]

A state-based soft-decision version of the estimator in (4) is possible, given the hidden Markov model of noisy speech. Ephraim[3] derived a soft-decision formulation of the MMSE estimate $\widehat{Y_t(k)}=E\{Y_t(k)|\mathbf{z}_0,\ldots,\mathbf{z}_{t+\tau_e-1},\lambda^z\}$ where the estimation delay in blocks, $\tau_e>0$, determines the number of future observations included and λ^z is the noisy speech model. Ephraim's implementation of this estimator can be shown[7] to be asymptotically equivalent to a weighted sum of MMSE DFT component estimators similar to (4), specifically,

$$\widehat{Y_t(k)}=\sum_{j=1}^{N_z}\sum_{m=1}^{M_z} H_{jm}(k)\times p(x_t=j,u_t=m|\mathbf{z}_0,\ldots,\mathbf{z}_{t+\tau_e-1},\lambda_z)Z_t(k), \tag{5}$$

where x_t is the model state, u_t is the Gaussian mixture, j is the noisy speech state index, m is the noisy speech mixture index, and

$$H_{jm}(k)=\frac{E\{|Y_t(k)|^2|x_t=j,u_t=m,\lambda^z\}}{E\{|Y_t(k)|^2|x_t=j,u_t=m,\lambda^z\}+E\{|V_t(k)|^2|x_t=j,u_t=m,\lambda^z\}} \tag{6}$$

is the approximate Wiener filter corresponding to noisy speech state j, mixture m. The state occupancy probabilities $p(x_t=j,u_t=m|\mathbf{z}_0,\ldots,\mathbf{z}_T,\lambda_z)$ are computed using the ''forward–backward'' procedure.[8] Since autoregressive (AR) HMMs[8] are used, a set of AR parameters is available for each state in both models. The second-order statistics required to evaluate (4) may be estimated using the AR parameters from the speech and noise models. One advantage of (5) over the classical Wiener filtering estimator (4) is that *a priori* knowledge of the dynamic behavior of the noise is utilized via the noisy speech HMM.

III. PROPOSED SYSTEM

The proposed speech enhancement system applies the HMM-based MMSE method to estimate sinusoidal synthesis parameters from noisy speech. From one point of view, this approach attempts to make the sinusoidal model more robust in the presence of noise by using a state-based estimator to find its parameters. From another point of view, this approach attempts to remove the residual noise associated with a HMM-based enhancement system by capitalizing on the noise suppression properties of the sinusoidal model. The harmonic spectral structure of voiced speech is implicitly emphasized and interharmonic noise is attenuated. A block diagram of the proposed system is shown in Fig. 1.

The lower half of Fig. 1 depicts the off-line training portion of the algorithm, i.e., selection of HMM and Wiener filter parameters. An ARHMM is trained using speech collected from a group of people representing potential users of the system. A similar model is trained with noise from the expected operating environment. A composite noisy speech model is then formed from the two independently trained models. Finally, an approximate Wiener filter is designed for every mixture of the noisy speech model.

The upper portion of Fig. 1 illustrates the operation of the algorithm after initial training. The box marked ''ACS'' represents calculation of the autocorrelation sequence (ACS) from a windowed block of noisy speech. Zero-padding is used in the implementation to achieve better frequency resolution for selection of the harmonic components of voiced speech. The forward–backward algorithm is applied to compute mixture likelihoods from the ACS for one or more noisy speech models. If more than one model is available, the model with the highest likelihood score is chosen. The harmonic widening factor ρ for the selected model is chosen empirically. The estimators associated with the noisy speech mixtures are weighted and summed to yield the soft-decision estimator or Wiener filter for the current block. This estimator is applied to the DFT components of the noisy speech to produce $\mathbf{Y}_t^{\text{hmm}}$. Voicing and fundamental frequency is estimated from $\mathbf{Y}_t^{\text{hmm}}$. If the current block of speech is classified as unvoiced or nonspeech, then the enhanced spectral components are not modified. If the block is voiced, the DFT components within $\pm\rho\omega_0/4\pi$ Hz of the harmonics are selected. All other components are set to zero. Finally, $\hat{\mathbf{y}}_t$ is reconstructed from the surviving spectral components using an inverse DFT, post-windowing, scaling, and the overlap-add procedure.

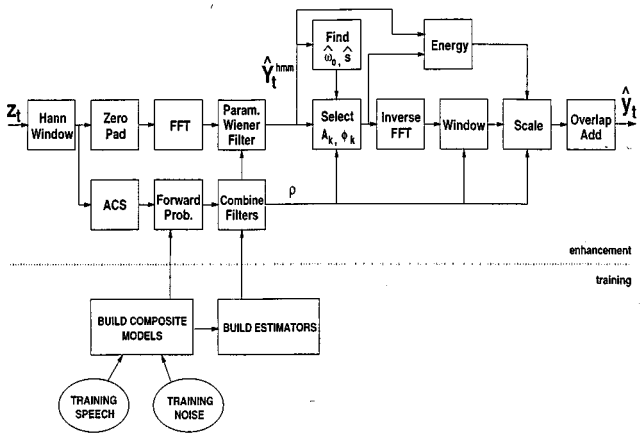


FIG. 1. Block diagram of proposed system.

A. Estimation of $\{A_\kappa, \phi_\kappa\}$

The system uses ergodic, continuous density, Gaussian mixture AR-HMMs. Two sets of training data are collected, one consisting of speech recorded under quiet conditions, and another consisting of noise from the expected operating environment. An AR-HMM representing clean speech and another representing noise are trained using the Baum–Welch reestimation procedure.[8] Both AR-HMMs have state output densities of the form

$$b_j(\mathbf{y})=\sum_m c_{jm}b_{jm}(\mathbf{y}), \tag{7}$$

where c_{jm} are mixture weights associated with the jth state output density and b_{jm} are Gaussian autoregressive densities.

The composite (noisy speech) model λ_z is constructed by using the assumptions of statistical independence and additivity of the speech and noise. The new state output densities are obtained by convolving each speech state output density with every noise state output density and approximating the result with an AR Gaussian density of sufficient order. The new transition probability matrix is $A_c=A_y\otimes A_v$, where $\otimes$ is the Kronecker product and A_y and A_v are the transition probability matrices of the clean speech and noise models, respectively. If the speech model has N_y states with M_y mixtures per state and the noise model has N_v states with M_v mixtures per state, then the composite model will have N_yN_v states with M_yM_v mixtures per state.

During enhancement, likelihoods are first computed for all mixtures of the noisy speech model using the forward

algorithm[8] ($\tau_e=1$). The score for the mth mixture in the jth state is

$$\varphi_{jm}(t)=\alpha_j(t)b_j^m(\mathbf{z}_t). \tag{8}$$

This is the likelihood that observation $\mathbf{z}_t$ was generated according to the mth mixture density of state j. In Eq. (8),

$$\alpha_j(t)=p(\mathbf{z}_0,\ldots,\mathbf{z}_t,x_t=j|\lambda_z) \tag{9}$$

is computed according to the forward algorithm,

$$b_j^m(\mathbf{z}_t)=\frac{c_{jm}b_{jm}(\mathbf{z}_t)}{b_j(\mathbf{z}_t)}, \tag{10}$$

and b_{jm} is the AR Gaussian density of the mth mixture in the jth noisy speech state. Spectral components are then estimated by applying a weighted combination of MMSE estimators to each DFT component of the input, i.e.,

$$\hat{Y}_{\mathrm{hmm}_t}(k)=\left[\sum_j\sum_m \varphi_{jm}(t)H_{jm}(k)\right]Z_t(k), \tag{11}$$

where

$$H_{jm}(k)=\frac{P'_{yy,jm}(k)}{P'_{yy,jm}(k)+\gamma_{jm}^{-1}P'_{vv,jm}(k)}. \tag{12}$$

Here, $H_{jm}(k)$ is an approximation of (4), the MMSE estimator of the kth DFT component, in terms of the AR parameters $(\sigma_y^2,\{a_n^y\},\sigma_v^2,\{a_n^v\})$ from the clean speech and noise HMMs corresponding to noisy speech state j and mixture m:

$$P_{yy}(k)=E\{|\widehat{Y_t(k)}|^2\}=\frac{\sigma_y^2}{|1+\Sigma_n a_n^y e^{jn\omega_k}|^2}, \tag{13}$$

$$P_{vv}(k)=E\{|\widehat{V_t(k)}|^2\}=\frac{\sigma_v^2}{|1+\Sigma_n a_n^v e^{jn\omega_k}|^2}. \tag{14}$$

The primes ($'$) in Eq. (12) indicate that $P_{yy}(k)$ and $P_{vv}(k)$ have been normalized to unit energy so that γ_{jm} in (12) is a signal-to-noise ratio. The value of γ_{jm} may be adjusted if the SNR is determined to be significantly different than the γ_{jm} obtained during training.

B. Estimation of ω_0

Streit and others have demonstrated the utility of hidden Markov models in frequency line tracking.[9] A conceptually similar approach is formulated here for fundamental frequency estimation of voiced speech. Assume that the fundamental frequency track is bounded below by $\omega_{0_{\min}}$ and bounded above by $\omega_{0_{\max}}$. Divide the interval $[\omega_{0_{\min}},\omega_{0_{\max}}]$ evenly into N^ω cells and denote the center frequency of each cell

$$\omega_{0(i)}=\left(i+\frac{1}{2}\right)\frac{\omega_{0_{\max}}-\omega_{0_{\min}}}{N^\omega},$$

where $i=0,\ldots,N^\omega-1$. If the quantized fundamental frequency is assumed to behave as a first-order, stationary Markov chain, a hidden Markov model may be defined for ω_0 by associating each cell with a state of the model. Unvoiced and nonspeech blocks may be assigned to an additional ''garbage'' state. The new model is denoted $\lambda^\omega=\{\pi^\omega,\mathbf{A}^\omega,B^\omega\}$. The state transition matrix $\mathbf{A}^\omega$ then characterizes the dynamic behavior of ω_0. The state of the model is observed through some parametrization of speech that strongly indicates its fundamental frequency. The state output observation sequence is denoted $O=(\mathbf{o}_0,\mathbf{o}_1,\ldots,\mathbf{o}_{t+\tau_\omega})$ where τ_ω is the fundamental frequency estimation delay in blocks. The density parameters must be selected using a *supervised* training procedure to maintain the association of the HMM states and frequency cells.

The maximum likelihood (ML) fundamental frequency trajectory given the model satisfies

$$X_{\mathrm{ML}}^\omega=\arg\max_{X^\omega} f(O,X^\omega|\lambda^\omega).$$

The Viterbi algorithm can be used to efficiently solve for the ML pitch track, $X_{\mathrm{ML}}^\omega=(\hat{x}_0^\omega,\ldots,\hat{x}_{t+\tau_\omega}^\omega)$.[10] The frequency corresponding to $\hat{x}_t^\omega$ is then chosen as the fundamental $\hat{\omega}_0$.

As in other hidden Markov modeling applications, proper choice of parameters $\mathbf{o}_t$ is critical to performance. A parametrization that strongly indicates pitch and is well suited to the proposed speech enhancement algorithm is desirable. The synthesis procedure for voiced speech used in the proposed algorithm involves sampling the set of estimated DFT components at the bins corresponding to harmonics. If the noise is completely removed by the Wiener filter and if the harmonics are perfectly resolved, then the unique harmonic frequency-sampling of $\hat{\mathbf{Y}}_t^{\mathrm{hmm}}$ that preserves its energy is that with fundamental ω_0. Although neither of these conditions can be met exactly, it is still possible to determine the fundamental frequency by finding the harmonic sampling that preserves *most* of the energy of a *flattened* version of $\hat{\mathbf{Y}}_t^{\mathrm{hmm}}$. Spectral flattening is necessary because of the effects of windowing on the voiced speech harmonics. Unless the spectrum is flattened prior to sampling, spectral leakage in the high-energy formant regions will bias the ω_0 estimate. If the flattened spectrum is denoted $\tilde{\mathbf{Y}}_t^{\mathrm{hmm}}$, then the parametrization based on spectral sampling is

$$\mathbf{o}_t=\left[\sum_{n\in\mathscr{W}_1}\left|\tilde{Y}_t^{\mathrm{hmm}}\left(\left\lfloor\frac{n\omega_{0(0)}\tilde{K}}{2\pi F_s}\right\rfloor\right)\right|^2\cdots\right.$$
$$\left.\sum_{n\in\mathscr{W}_{N^\omega}}\left|\tilde{Y}_t^{\mathrm{hmm}}\left(\left\lfloor\frac{n\omega_{0(N^\omega-1)}\tilde{K}}{2\pi F_s}\right\rfloor\right)\right|^2\right]^T, \tag{15}$$

where F_s is the sampling frequency, $\mathscr{W}_m\in\mathbb{Z}^{N_{\mathrm{harm}}}$, and N_{harm} is the number of harmonics included in the sum.

Although it was developed independently, the procedure for computing $\mathbf{o}_t$ can be considered a refinement of the pitch extraction method of Steiglitz *et al.* that fits a finite Fourier series to the speech waveform in the least squares sense.[11] McAulay and Quatieri[12] also proposed a procedure based on sampling the spectral magnitudes. However, in that case the objective was to find the fundamental frequency that minimized the mean squared error between speech synthesized with a peak-picked spectrum and speech synthesized with harmonic components. Their method used a time-varying

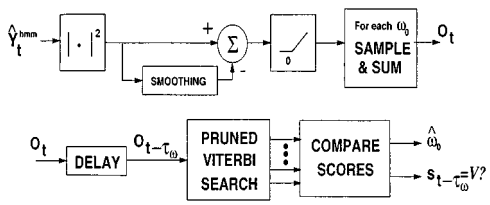


FIG. 2. Block diagram of simplified fundamental frequency estimation algorithm.

analysis window whose size is based on the average pitch. The proposed method is simpler in that it does not require a preliminary estimate of the average pitch.

The upper half of Fig. 2 illustrates the calculation of $\mathbf{o}_t$. First, the squared magnitudes of the estimated DFT components $\hat{\mathbf{Y}}^{\text{hmm}}$ are computed. Spectral flattening is accomplished by subtracting a smoothed version of the magnitude squared spectrum and setting any negative components to zero. Smoothing is performed using a moving average filter. Entries of $\mathbf{o}_t$ are computed by summing the harmonic samples of the flattened squared-magnitude spectrum for each of the fundamental frequency candidates. The set $\mathscr{W}_m$ is defined to contain the indices of the largest N_{harm} consecutive harmonic samples for the mth fundamental frequency candidate where $N_{\text{harm}}=\lfloor \pi F_s/\omega_{0_{\max}} \rfloor$.

If the HMM-based fundamental frequency estimation algorithm was implemented directly, the cost in terms of computational load and memory would be high. Since the largest cell width that provides satisfactory results for the proposed speech enhancement algorithm is about 1 Hz and the frequency range of interest is approximately [70,400] Hz a 332-state model is needed! Therefore, several simplifications are made to avoid an unnecessarily high computational load. First, each state output probability density b_j^{ω} is replaced by

$$\hat{b}_j^{\omega}(\mathbf{o}_t)=\frac{o_t(j)}{\|\hat{\mathbf{Y}}^{\text{hmm}}\|_2}$$

so that explicit calculation of state output probabilities is avoided. In addition, the state transition probabilities for the quantized fundamental frequency computed based on a distribution of continuous, relative incremental ω_0 which is approximated as Gaussian, i.e.,

$$\hat{a}_{ij}^{\omega}=\int_{q_1}^{q_2}\frac{1}{\sqrt{\pi\sigma^{\omega 2}}}\,e^{-q^2/2(\sigma^{\omega})^2}\,dq, \tag{16}$$

where $q_1=\min|\omega'-\omega''|/\omega'$ and $q_2=\max|\omega'-\omega''|/\omega'$ for $\omega' \in$ cell i and $\omega'' \in$ cell j, $q=|\omega_{0(i)}-\omega_{0(j)}|/\omega_{0(i)}$, and σ_ω^2 is determined experimentally. Although a histogram of the relative change measured from $\mathbf{o}_t$ is asymmetric and multimodal, the main lobe is centered on ω_0, is significantly larger than the others, and is fit reasonably well by a normal curve. Equation (16) is evaluated using a crude rectangle-rule approximation. Since $b_j^{\omega}(\mathbf{o}_t)$ and a_{ij} are computed online, there is no need for training in the usual sense. Finally, the Viterbi search is ''pruned'' such that only the N_{tracks} largest values of $b_j^{\omega}(\mathbf{o}_t)$ are considered. One block of delay is introduced ($\tau_\omega=1$) and the fundamental frequency candidate contained in the most likely ω_0 trajectory is chosen. Instead of using a garbage state to determine voicing, the likelihood of the ML ω_0 trajectory is compared to that of the other trajectories considered. If the sum of the absolute differences exceeds a threshold, the block is classified as voiced.

C. Selection of harmonic components

The frequency-domain output of the HMM-based algorithm (denoted $\hat{\mathbf{Y}}_{\text{hmm}}$) is used to determine the parameters for sinusoidal synthesis. Once an estimate of the fundamental frequency, $\widehat{\omega_0}$, has been determined, the amplitudes of the harmonic components are chosen to be $\hat{A}_{t,\kappa}=2|\hat{Y}_{\text{hmm}_t}(k)|$, where k is the index of the component closest in frequency to $\kappa\widehat{\omega_0}$. The phases for the harmonic model are $\hat{\phi}_{t,\kappa}=\angle\hat{Y}_{\text{hmm}_t}(t)$. This is equivalent to using the measured phase, $\angle Z_t(k)$, since φ_{jm} and H_{jm} are real. If the block of speech is unvoiced, all components are used for synthesis.

Given $\widehat{\omega_0}$, $\hat{A}_{t,\kappa}$, and $\hat{\phi}_{t,\kappa}$, the parameter ρ determines the number of additional DFT components selected during voiced speech. The best values of ρ were chosen experimentally as a function of average input SNR and are shown in Fig. 3. Figure 4 illustrates the effect of the choice of the harmonic widening factor ρ on the average output SNR [see Eq. (18)] computed using utterances from sixty speakers at 0 dB input SNR (details of the experiments are given in Sec. V). The average output SNR using the HMM-based estimator is shown for comparison. When $\rho=1$, the proposed algorithm is equivalent to the ''baseline'' HMM-based algorithm. When ρ is fixed at zero for all voiced segments, a purely harmonic model is used for voiced speech. This results in an increase of 0.4 dB in average output SNR since the noise between the chosen components has been completely suppressed. However, errors in the fundamental frequency estimate cause a slight degradation in this case. In addition, the fact that voiced speech is not always perfectly periodic means that some important components are missed. When $\rho=0.4$ the best compromise between fundamental frequency estimator/modeling errors and interharmonic noise suppression is achieved. It should be noted that the average output SNR is not very sensitive to choice of ρ. Values between 0.3, and 0.6 give similar results.

The vector of selected components is denoted $\hat{\mathbf{Y}}_{\text{sm}}$. After sinusoidal synthesis, the gain of the reconstructed signal is adjusted by multiplying with

$$g=\frac{\|\hat{Y}_{\text{hmm}}\|_2}{\|\hat{\mathbf{Y}}_{\text{sm}}\|_2} \tag{17}$$

to produce the clean speech estimate, $\hat{\mathbf{y}}_t$.

IV. SIMULATION RESULTS

The proposed speech enhancement system was evaluated using speech from the TIMIT database.[13] One hundred sentences (approximately 5 min total) were selected for training the speech model. Fifty were chosen from the ''si'' natural phonetic sentences (25 male, 25 female) and fifty were chosen from the ''sx'' phonetically compact sentences (25

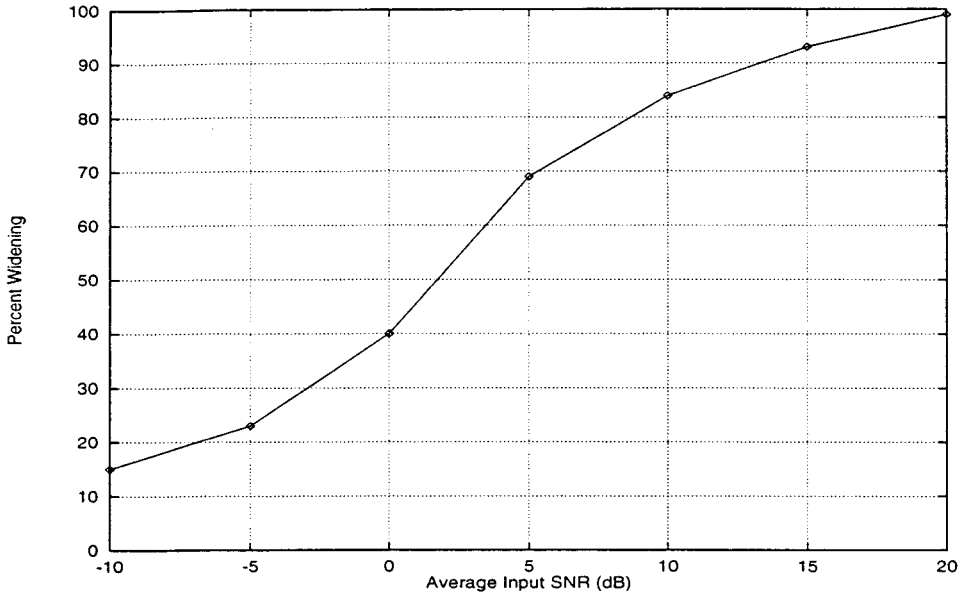


FIG. 3. Best widening factor versus average input SNR.

male, 25 female). Speech was modeled by an eight-state, four-mixture, 12th-order AR-HMM. ''Speech-shaped'' noise from the NOISEX-92 database[14] was used to train the noise model. Speech-shaped noise is white noise that has been spectrally shaped using a filter whose frequency response matches the long-term spectrum of speech. Spectrally shaped noise was judged to be a better choice than white noise for demonstrating the merits of the proposed algorithm since it produces a lower average SNR in the first and second formant regions of voiced speech. Noise was modeled by a one-state, single mixture, fourth-order AR-HMM. The composite model was an eight-state, four-mixture, 16th-order AR-HMM. The algorithm was tested using 60 sentences from the TIMIT database. Thirty were taken from the ''si'' sentences (15 male, 15 female) and thirty were taken from the ''sx'' sentences (15 male, 15 female). Speech-shaped noise was then added to the test sentences at SNRs between -10 and 20 dB in increments of 5 dB. Training and testing were carried out at these input SNRs. The resulting output SNRs averaged over the entire testing set are shown in Fig. 5. Here, SNR is defined as

$$\mathrm{SNR}=10\log_{10}\frac{\Sigma_n y^2(n)}{\Sigma_n(y(n)-\hat{y}(n))^2}, \tag{18}$$

where the sums are taken over all nonoverlapping blocks of clean speech whose energy is at least -40 dB. The average output SNR obtained when using the fundamental frequency

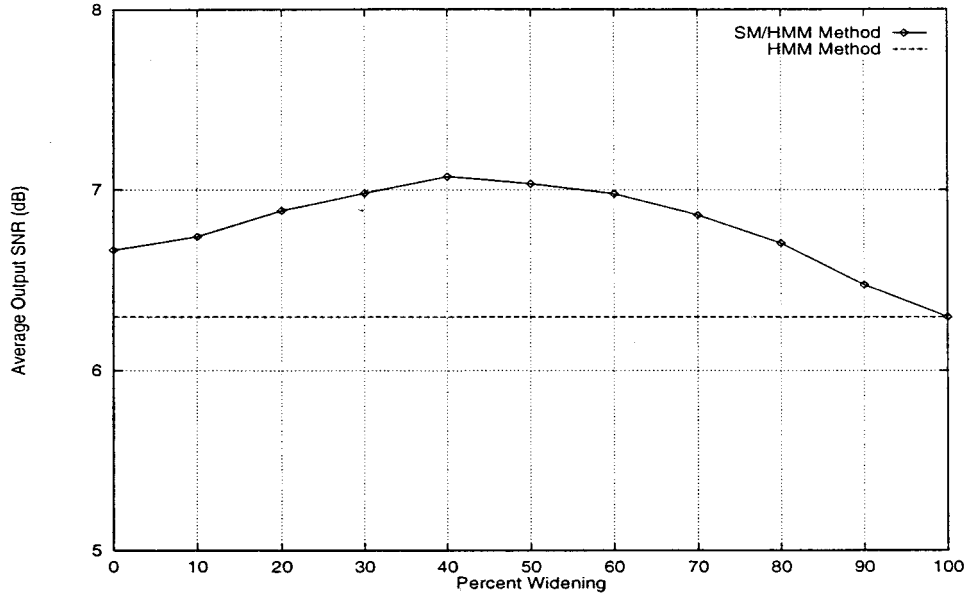


FIG. 4. Effect of choice of $\rho(\times 100\%)$ on average output SNR at 0 dB input SNR.

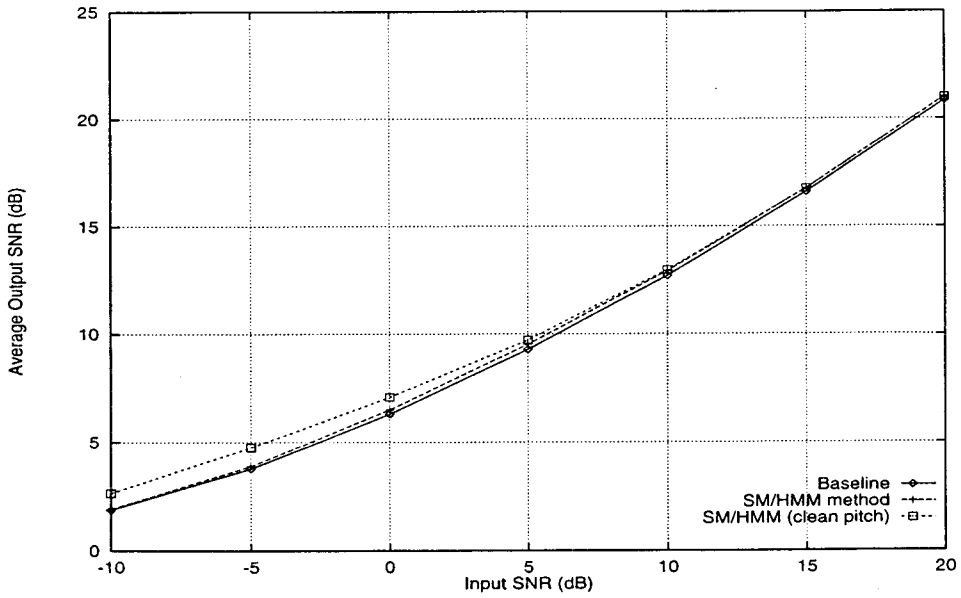


FIG. 5. Comparison of baseline HMM-based algorithm, SM/HMM-based algorithm, and SM/HMM-based algorithm with ω_0 estimated from clean speech.

estimated from clean speech is also shown in Fig. 5 for comparison. For input SNRs at and below 0 dB, fundamental frequency estimation errors prevent the proposed algorithm from achieving a large SNR improvement.

The effect of fundamental frequency errors was studied by adding uniformly distributed random noise to the values of ω_0 estimated from clean speech. The results are shown in Fig. 6 for an input SNR of 0 dB. When the maximum error $\widehat{\omega_0}$ exceeds 2.5 Hz, the benefit of the proposed algorithm is lost. However, the value of ρ may be adjusted to compensate for fundamental frequency estimator errors. The average output SNR using the best ρ (chosen experimentally) is plotted in Fig. 6 to illustrate this.

Although the improvement in average output SNR is small, the subjective improvement is noticeable, particularly for female speakers. Figure 7 shows the segmental SNR for an utterance by a female speaker. The text of the sentence is shown at the top of the plot. The segmental SNR is computed using (18) for each block rather than the entire utterance. As much as 2 dB improvement is shown in Fig. 7 during voiced segments. In particular, /æ/, /e/, /u/, and /l/ are most improved. The degradation in segmental SNR during /s/ is due to erroneous voicing decisions. The degradation is small relative to the overall improvement of /s/ and is not audibly perceptible.

Although the HMM-based speech enhancement algorithm does not produce tonal artifacts in the output speech as some other methods (e.g., spectral subtraction[15]) do, it is

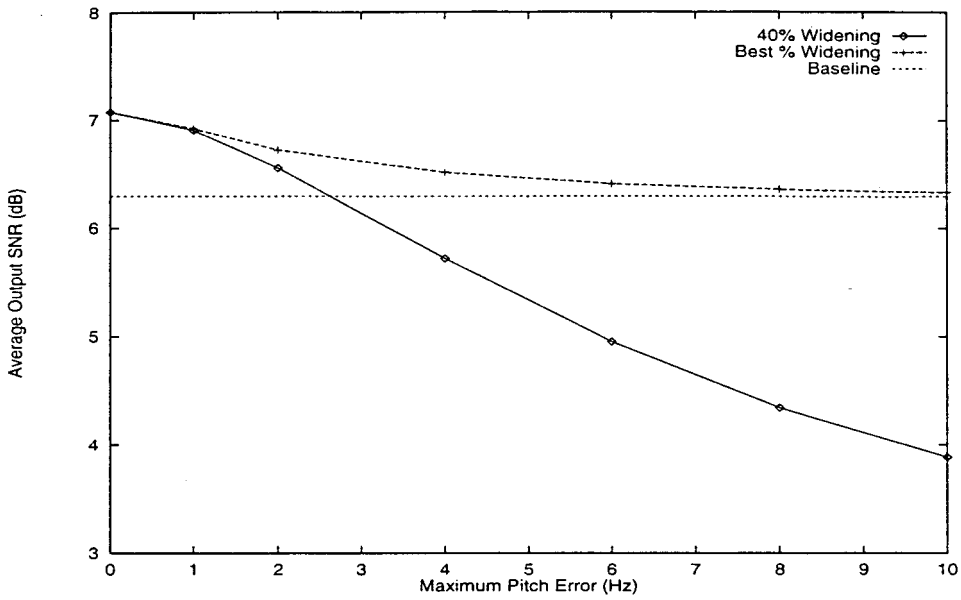


FIG. 6. Effect of errors in $\widehat{\omega_0}$ at 0 dB input SNR. Output SNR obtained using the baseline HMM-based method[3] is shown for comparison.

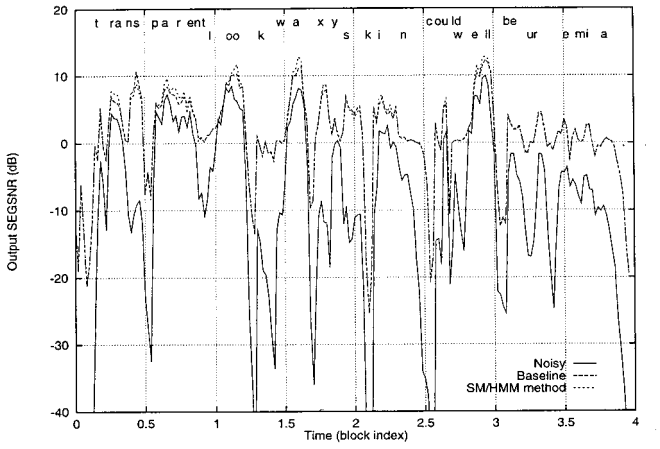


FIG. 7. Segmental SNR at 0 dB total input SNR for a female speaker. Segmental SNR obtained using the baseline HMM-based method[3] is shown for comparison.

known to produce ''a low-level structured residual noise'' which is more noticeable for female speakers.[3] The residual noise increases in intensity as the SNR is reduced. When the input SNR is at or below 0 dB, processed speech sounds ''hoarse'' or ''breathy.'' The proposed algorithm noticeably attenuates this residual noise. In fact, the best noise reduction is obtained for high-pitched speakers since the harmonic components are spaced further apart.

Formal intelligibility and quality tests have not been carried out at the time of this writing. However, casual tests involving four listeners and the 60 processed TIMIT sentences indicate that (a) the low-level residual noise associated with HMM-based algorithms was perceptibly reduced, (b) the intelligibility of the processed speech was not affected, and (c) there were no additional artifacts (e.g., vocoding noise) introduced by harmonic modeling.

V. CONCLUSIONS

A new HMM-based speech enhancement scheme has been presented. Noise reduction is achieved by applying the HMM-based MMSE estimator to find the harmonic sinusoidal model parameters of clean speech from speech corrupted by additive noise. The proposed technique was compared to the standard HMM-based approach with respect to average increase in total SNR, segmental SNR and subjective quality. A small gain in average output SNR was obtained. At 0 dB input SNR, the proposed algorithm improved the segmental SNR during voiced speech by as much as 2 dB. Informal listening tests indicate that the low-level residual noise associated with HMM-based algorithms was perceptibly reduced.

ACKNOWLEDGMENTS

This work has been supported by the Mobile Research Council of Intel Corporation. The authors would like to thank Brian Mears of Intel Corporation for his comments and suggestions.

[1] R. J. MacAulay and T. F. Quatieri, ''Low-Rate Speech Coding Based on the Sinusoidal Model,'' in *Advances in Speech Signal Processing*, edited by S. Furui and M. M. Sondhi (Dekker, New York, 1992), pp. 165–208.
[2] T. F. Quatieri and R. G. Danisewicz, ''An Approach to Co-Channel Talker Interference Suppression Using a Sinusoidal Model for Speech,'' IEEE Trans. Acoust. Speech Signal Process. **38**, 56–59 (1990).
[3] Y. Ephraim, ''A Minimum Mean Square Error Approach for Speech Enhancement,'' Proc. ICASSP-90, May 1990, pp. 829–832.
[4] J. M. Tribolet and R. E. Crochiere, ''Frequency Domain Coding of Speech,'' IEEE Trans. Acoust. Speech Signal Process. **27**, 512–530 (1979).
[5] T. F. Quatieri and R. J. McAulay, ''Noise Reduction Using a Soft-Decision Sine-Wave Vector Quantizer,'' Proc. ICASSP-90, May 1990, pp. 821–824.
[6] J. S. Lim and A. V. Oppenheim, ''Enhancement and Bandwidth Compression of Noisy Speech,'' Proc. IEEE **67**, 1586–1604 (1979).
[7] M. E. Deisher, ''State-Based Noise Reduction Using the Sinusoidal Speech Model,'' Ph.D. dissertation, Arizona State University, May 1996.
[8] L. R. Rabiner, ''A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition,'' Proc. IEEE **77**, 257–286 (1989).
[9] R. L. Streit and R. F. Barrett, ''Frequency Line Tracking Using Hidden Markov Models,'' IEEE Trans. Acoust. Speech Signal Process. **38**, 586–598 (1990).
[10] G. D. Forney, Jr., ''The Viterbi Algorithm,'' Proc. IEEE **61**, 268–277 (1973).
[11] K. Steiglitz, G. Winham, and J. Petzinger, ''Pitch Extraction by Trigonometric Curve Fitting,'' IEEE Trans. Acoust. Speech Signal Process. **23**, 321–323 (1975).
[12] R. J. McAulay and T. F. Quatieri, ''Pitch Estimation and Voicing Detection Based on a Sinusoidal Speech Model,'' Proc. ICASSP-90, May 1990, pp. 249–252.
[13] J. S. Garofolo, ''Getting Started with the DARPA TIMIT CD-ROM: An Acoustic Phonetic Continuous Speech Database,'' National Institute of Standards and Technology, December 1988.
[14] A. Varga *et al.*, ''The NOISEX-92 Study on the Effect of Additive Noise on Automatic Speech Recognition,'' DRA Speech Research Unit (UK) and TNO Institute for Perception (Netherlands), June 1992.
[15] S. F. Boll, ''Suppression of Acoustic Noise in Speech Using Spectral Subtraction,'' IEEE Trans. Acoust. Speech Signal Process. **27**, 113–120 (1979).
[16] M. R. Schroeder, ''Period histogram and product spectrum: New methods for fundamental frequency measurement,'' J. Acoust. Soc. Am. **43**, 829–834 (1968).

Incorporating phonetic properties in hidden Markov models for speech recognition

Ramachandrula N. V. Sitaram[a] and Thippur Sreenivas
Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore-560012, India

(Received 6 March 1996; revised 8 October 1996; accepted 14 March 1997)

In this paper the incorporation of important phonetic properties into hidden Markov models (HMM) is studied. Phones have characteristic properties such as unique temporal structure, context sensitive behavior and specific duration, etc. New HMMs which incorporate the above properties with additional degrees of freedom to the standard HMM states are proposed. The use of each of the phonetic property for speech recognition is demonstrated using the new HMMs. All the algorithms required for using these new models in various applications of speech recognition have been presented. Experimental comparison with the standard discrete HMM for a speaker-independent continuous speech phone recognition task show that consistent improvement is achieved by the new models. © *1997 Acoustical Society of America.* [S0001-4966(97)02007-9]

PACS numbers: 43.72.Ne [JS]

INTRODUCTION

Acoustic phonetic decoding is one of the classic speech recognition problems that has been addressed for more than three decades with gradual success. Better success in this can change the methodology of many current speech recognition systems that are based on word detection. Since a phone is the smallest unit of speech, any improvement in its recognition will lead to improvement in word recognition. The acoustic manifestations of phones have specific properties, such as unique temporal structure, duration, context sensitive behavior, etc. Most successful phone recognizers try to use as many of these properties as possible (Lee and Hon, 1989; Ljolje and Levinson, 1991; Robinson and Fallside, 1991; Lamel and Gauvain, 1993; Ljolje, 1994). In this paper, we study the incorporation of these properties directly into an HMM, resulting in new HMMs with more degrees of freedom and better performance.

Spectrograms reveal the dynamic behavior of a phone, particularly stop consonants and diphthongs, which have a unique temporal structure that differentiates them from other phones. The phones also have significant variability due to different talkers and different contexts, which has to be specifically modeled. Expert spectrogram readers identify phones based on the dynamic behavior of formant tracks rather than their absolute location. Therefore, proper statistical modeling of the dynamic behavior of phones and its incorporation into phone recognizers can achieve better results.

Among the differentiating features of a phone, duration is an important cue. The vowels /ae/ (e.g., b*a*t) and /eh/ (e.g., b*e*t) are acoustically similar but their durations are very different. While discriminating between confusable words, like *rider* and *writer*, where the differentiating phones /d/ and /t/ are very similar, the duration of these phones plays an important role. Due to talker variability, the spectral structure of some phones may become closer to different phones, but the duration offers an important cue for discrimination.

In realizing a sound, the articulators are highly influenced by the articulatory configuration of the preceding and following sounds. Therefore there is a systematic variability in the acoustic manifestation of a sound due to its context. The context affects the spectral content, temporal structure, and also the duration of a phone. In speech recognition, to counter this type of variability, knowledge of different contextual effects on each phone has to be used.

Most of the present day speech recognizers are based on hidden Markov modeling (HMM), which is a statistical tool to model dynamic patterns. Typically, a spoken word is modeled by an HMM of ''*N*'' states, where each state is intended to characterize a single phone. But this single state to phone relationship is not well realized in practice because a single observation probability density function (pdf) of an HMM state cannot model the dynamic behavior of the phone; such a stationary model, at best, suits a stationary segment of a phone. Figure 1(a) shows spectrograms of two nonstationary phones, diphthongs /ay/ (e.g., b*i*te) and /aw/ (e.g., b*ou*t). These two phones have distinct temporal structure such as rising second formant vs falling second formant. But single discrete pdfs (vector quantized) of each diphthong, estimated using many occurrences of them [Fig. 1(b)] are very similar, thus failing to capture differences in the temporal structure, leading to recognition errors. One commonly followed method to ease the above problem is to use constrained left to right HMMs of multiple states, per phone, so that each state is modeling a short stationary segment of the phone, and the left–right constraint preserves the sequencing (temporal structure) of subphonetic segments.

However, even using multiple states to model a phone is inadequate in discriminating confusable words like *rider* and *writer*, where the duration of the differentiating phones is the important discriminating feature. For example, if three states are catered to each phone in these word models, there is a need to model the total duration spent in the three states of each phone and use this duration probability in the forward computation of word utterance likelihood. However, specific duration modeling by multiple connected states in HMM is

[a]Electronic mail: sitaram@protocol.ece.iisc.ernet.in

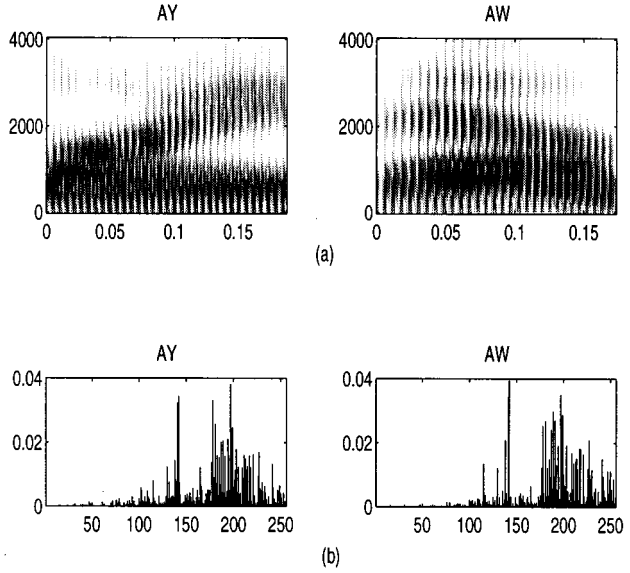


FIG. 1. (a) Typical spectrograms of diphthongs /ay/ and /aw/. (b) Discrete pdfs (256 size codebook) of two diphthongs, /ay/ and /aw/, estimated from the TIMIT database.

not achieved automatically in forward–backward training. It is possible to add a penalty for the duration spent in each state at the end of Viterbi scoring of the word. There are several attempts in the literature (Ramesh and Wilpon, 1992; Rabiner, 1989; Levinson, 1986; Russel and Moore, 1985; Russel and Cook, 1987) to model the state duration in a HMM. In particular, for modeling phone duration to discriminate confusable words, Russel and Cook (1987) have used semi-HMMs. Since a HMM state occupancy follows a geometric distribution, whereas the required phone duration pdf is a Gamma distribution (Crystal and House, 1982), they have used a HMM without self-transitions, for explicitly modeling the state duration. They have reported improved performance using semi-HMMs for discriminating confusable words. Thus even though a single HMM state can be made to model duration of phones accurately, it cannot model the temporal structure as mentioned earlier.

One important assumption of an HMM is that the observation vectors corresponding to consecutive speech frames are independent of each other. Such a model simplifies the difficulty of evaluating the probability of long sequence of speech frames, in addition to representing the acoustic variability in a phone due to many talkers. However, in a specific phone, successive frames have a certain time-dependent pattern which differentiates it from other phones. Also, the initial and final portions of a phone are affected by the neighboring phones. Triphone models (Lee *et al.*, 1990) have been developed to accommodate this coarticulation effect. However, triphone models also use standard HMMs, and hence have the weakness of not capturing the temporal dependence in speech.

We would like to incorporate both the temporal structure and durational property of a phone into an HMM state, which have rather contradicting requirements; i.e., temporal modeling requires more states to be catered to a phone and duration modeling requires a single state to be catered to a phone. If we can make the HMM state more powerful so that it can model temporal structure of the phone, then durational constraint can also be incorporated. Clearly, there is a need to develop new HMMs which can better model the dynamics of speech, its characteristic durations, and the contextual effects. We would then have realized a one-to-one relationship between a phone and an HMM state. Also, it would resolve the often faced problem of choosing the number of states in an HMM for a given word. In this study, three new HMMs are proposed, each incorporating an important property of the phone with the addition of a new degree of freedom to the HMM state.

The proposed new models have been tested for a phone recognition task. The single state per phone feature allows the phone recognizer to be an ergodic network with the number of states equal to the number of phones. The HMM parameters are derived using the phonetically labeled speech database, TIMIT. Phonetic decoding of a test utterance is obtained through maximum likelihood state sequence decoding. The merit of the new models is assessed using relative improvement of phone recognition scores; however, no effort is made to improve the absolute phone recognition scores to that of the state of the art. This task has been deferred because of the complexity involved in repeating the experiments for different acoustic front ends and use of continuous observation density HMMs. In all experiments, discrete HMMs with a single codebook are used for simplicity of training.

Section I discusses how a single HMM state can be made to model the temporal structure of a phone. In this section a new HMM with a restriction on the probabilities of observation symbols emitted from the state, based on the duration spent in the state is proposed, along with the solutions of the three standard problems of HMM. In Sec. II, HMM state duration modeling is discussed with a special reference to inhomogeneous HMM (Ramesh and Wilpon, 1992) for capturing phone durations. And Sec. III deals with coarticulation modeling, with the proposal of a new HMM with context-dependent observation pdfs, followed by Sec. IV, comprised of experiments and discussions.

I. TEMPORAL STRUCTURE MODELING

The acoustic manifestation of phones, such as long vowels, is quasistationary over most of the segment, and hence they can be modeled using a single observation pdf of a state. However, nonstationary phones, such as stop consonants, diphthongs, etc., are characterized by their dynamic behavior and they may not have any quasistationary region. The specific sequence of acoustically distinct events characterize these phones, in spite of the acoustic variability within a subphonetic segment. A different sequencing of the same acoustic subsegments would correspond to a different phone or phones. Thus recognition of a phone requires recognizing the order of events. If a single HMM state could realize such ordering of events, then it can characterize a nonstationary phone. This can be achieved by restricting the pdf of observation symbols emitted from a state to depend on the time spent in the state. Thus a set of time-dependent observation

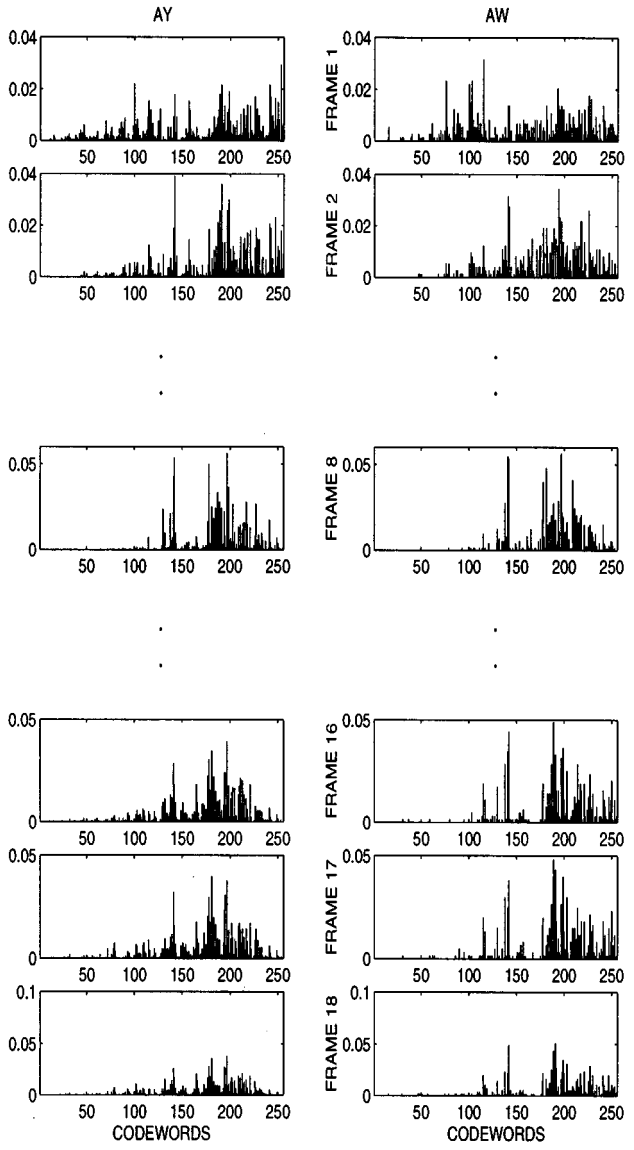


FIG. 2. Time-dependent discrete pdfs of diphthongs /ay/ and /aw/ shown labeled for each time frame.

symbol pdfs are associated with each state, each pdf corresponding to the time elapsed in the state. Considering discrete HMMs, the observation symbol distribution, $b_j(k)$, is made dependent on the duration d spent in the state j; i.e., it becomes $b_j(k,d)$, a different observation distribution after staying in state j for d clock intervals. Thus each state will have a sequence of observation distributions, one for each clock interval, until a maximum limit of D. For $d>D$, the distributions are truncated to $b_j(k,D)$; i.e., $b_j(k,d)=b_j(k,D)$ for $d>D$. The value of D has to be chosen so as to cover the maximum duration of the phone modeled by the state. This new extension is referred to as Trend-HMM (T-HMM).

Considering again the example of Fig. 1, one can see that the two sequences of discrete pdfs of both diphthongs /ay/ and /aw/ (see Fig. 2), calculated over many occurrences of the two phones at each frame instant are quite distinct at the frame numbers 16, 17, 18, etc., thus helping in their discrimination; whereas the combined single pdfs of Fig. 1(b) calculated from the whole length of phones are not distinct, causing confusion. Thus a T-HMM state which has a sequence of observation pdfs, one for every time frame, can model the temporal structure of a phone.

A. Training/scoring of T-HMM

Given the form of T-HMM, for it to be useful in speech applications, the re-estimation formulae for all the parameters have been derived. Also, the Viterbi algorithm for maximum likelihood state sequence decoding, required for the phone recognizer, is derived. First, the forward–backward algorithm, which is useful in the re-estimation of the parameters and also in the evaluation of an observation sequence given a model, is presented.

1. Forward–backward algorithm

Let two variables, $\alpha_t(i,d)$ and $\beta_t(i,d)$, denote the forward and backward probabilities used in the efficient computation of probability of the observation sequence given a T-HMM, λ. The forward variable $\alpha_t(i,d)$ is the joint probability of the partial observation sequence $\{O_1,O_2,\ldots,O_t\}$ such that the state occupied at time t is i and that the machine is in state i for the past $d-1$ instants, given the model λ; i.e.,

$$\alpha_t(i,d)=P(O_1,O_2,\ldots,O_t,q_t=i,d_t(i)=d|\lambda), \tag{1}$$

where $d_t(i)$ is the duration spent in state i at time t; here $\lambda=(\mathbf{A},\mathbf{B},\mathbf{\Pi},N,M)$, $\mathbf{A}$, $\mathbf{\Pi}$, M, N, have the same meaning as in a standard HMM (Rabiner, 1989); whereas, $\mathbf{B}$ denotes the set of duration-dependent observation probability distributions $b_j(k,d)$, $1\leqslant j\leqslant N$; $1\leqslant k\leqslant M$; $1\leqslant d\leqslant D$, with the statistical condition $\Sigma_{k=1}^{M}b_j(k,d)=1\forall d$ and j. Let $\mathbf{C}=\{v_k,1\leqslant k\leqslant M\}$ be the codebook of observation symbols. The forward probabilities are computed using the following algorithm:

Initialization: for $1\leqslant i\leqslant N$

$$\alpha_1(i,1)=\pi_i b_i(O_1,1), \tag{2}$$

$$\alpha_1(i,d)=0, \quad d>1; \tag{3}$$

Recursion:

$$\alpha_{t+1}(j,1)=\sum_{\tau=1}^{t}\sum_{i=1}^{N}\alpha_t(i,\tau)a_{ij}b_j(O_{t+1},1), \quad i\neq j \tag{4}$$

$$\alpha_{t+1}(j,d+1)=\alpha_t(j,d)a_{jj}b_j(O_{t+1},d+1); \tag{5}$$

$d=1,2,\ldots,t$, $j=1,2,\ldots,N$ and $t=1,2,\ldots,T-1$.

The forward variable $\alpha_{t+1}(j,1)$ indicates that a unit duration is spent in state j at $t+1$, and for this a non-self-state transition should take place into state j at t. Therefore the variable $\alpha_{t+1}(j,1)$ is evaluated by the summation of the probabilities of the partial observation sequence $\{O_1,O_2,\ldots,O_t,O_{t+1}\}$ emitted by all possible state sequences $\{q_1,q_2,\ldots,q_t,q_{t+1}\}$ with a requirement that at $t+1$ the state q_{t+1} is j, which could have been entered from any state, $q_t\neq j$. The variable $\alpha_{t+1}(j,d+1)$ in Eq. (5) could

be evaluated for various values of d from $\alpha_t(j,d)$ evaluated in previous recursions. Now the probability of the observation sequence is given by:

$$P[O|\lambda]=\sum_{d=1}^{T}\sum_{i=1}^{N}\alpha_T(i,d). \tag{6}$$

Similarly the backward variable $\beta_t(j,d)$ is defined as the probability of the partial observation sequence $\{O_{t+1},O_{t+2},\ldots,O_T\}$ with the condition that the state occupied at time t is i for duration $d_t(i)=d$, given the model λ, i.e.,

$$\beta_t(i,d)=P(O_{t+1},\ldots,O_T|q_t=i,d_t(i)=d,\lambda). \tag{7}$$

This variable required in re-estimation is calculated as follows:

Initialization: for $1\leqslant i\leqslant N$; and $1\leqslant d\leqslant T$,

$$\beta_T(i,d)=1; \tag{8}$$

Recursion: for $1\leqslant j\leqslant N$; $t=T-1$, $T-2,\ldots,1$; and $d=1,2,\ldots,t$;

$$\beta_t(j,d)=\left[\sum_{i=1,i\neq j}^{N}\beta_{t+1}(i,1)a_{ji}b_i(O_{t+1},1)\right]+\beta_{t+1}(j,d+1)a_{jj}b_j(O_{t+1},d+1). \tag{9}$$

The variables $\beta_t(j,d)$ are calculated $\forall t,j,d$, using the above equation iteratively, starting from $t=T-1$ backward. In the equation, the two terms correspond to all non-self-transitions and a self-transition from state j at t, respectively.

2. Re-estimation

Given an initial T-HMM, λ, and an observation sequence $\{O_1,O_2,\ldots,O_T\}$, the parameters of the T-HMM are re-estimated as given below. Two new variables, $\gamma_t(i,d)$ and $\xi_t(i,j,d)$, are defined and used in re-estimation. Defining

$$\gamma_t(i,d)=P(q_t=i,d_t(i)=d|O,\lambda), \tag{10}$$

$$\Rightarrow\gamma_t(i,d)=\frac{\alpha_t(i,d)\beta_t(i,d)}{P(O/\lambda)},\quad \forall i,t \text{ and } 1\leqslant d\leqslant t. \tag{11}$$

Now the duration-dependent observation probabilities **B** are re-estimated as follows: For $1\leqslant i\leqslant N$; $1\leqslant k\leqslant M$:

$$b_i(k,d)=\frac{\Sigma_{t=1}^{T}\gamma_t(i,d) \text{ such that } O_t=v_k}{\Sigma_{t=1}^{T}\gamma_t(i,d)},\quad 1\leqslant d<D \tag{12}$$

$$b_i(k,D)=\frac{\Sigma_{t=1}^{T}\Sigma_{d=D}^{T}\gamma_t(i,d) \text{ such that } O_t=v_k}{\Sigma_{t=1}^{T}\Sigma_{d=D}^{T}\gamma_t(i,d)}. \tag{13}$$

Here the observation symbol v_k corresponds to the codeword k. Defining

$$\xi_t(i,j,d)=P(q_t=i,q_{t+1}=j,d_t(i)=d|O,\lambda) \tag{14}$$

for $1\leqslant i,j\leqslant N$; and $1\leqslant d\leqslant T$;

$$\xi_{t\ i\neq j}(i,j,d)=\frac{\alpha_t(i,d)a_{ij}b_j(O_{t+1},1)\beta_{t+1}(j,1)}{P(O/\lambda)}, \tag{15}$$

$$\xi_{t\ i=j}(i,i,d)=\frac{\alpha_t(i,d)a_{ii}b_i(O_{t+1},d+1)\beta_{t+1}(i,d+1)}{P(O/\lambda)}. \tag{16}$$

The transition probabilities **A** are re-estimated as follows:

$$a_{ij}=\sum_{t=1}^{T-1}\sum_{d=1}^{t}\xi_t(i,j,d)\Big/\sum_{t=1}^{T-1}\sum_{d=1}^{t}\gamma_t(i,d),\quad 1\leqslant i,j\leqslant N. \tag{17}$$

The initial state occupancy probabilities **Π** are re-estimated as

$$\pi_i=\gamma_1(i,1),\quad 1\leqslant i\leqslant N. \tag{18}$$

3. Viterbi algorithm

Given an observation sequence $\{O_1,O_2,O_3,\ldots,O_T\}$ and a T-HMM, λ, the problem is to determine the hidden state sequence which contributes maximum probability. This problem occurs in all CWR and CSR applications and it can be solved as follows.

Let $\delta_t(i,d)$ be the probability of the best state sequence $\{q_1,q_2,\ldots,q_t\}$ for the partial observation sequence $O^t=O_1,O_2,\ldots,O_t$, with q_t being i and duration spent in state i, at t, being d, given λ, thus

$$\delta_t(i,d)=\max_q\{P[q_1,\ldots,q_{t-1},q_t=i,d_t(i)=d,O^t|\lambda]\}, \tag{19}$$

where $d_t(i)$ is the duration spent in state i at time t for the best path ending at $q_t=i$. The $\delta_t(i,d)$ are computed as shown below.

Initialization: for $1\leqslant i\leqslant N$

$$\delta_1(i,1)=\pi_i b_i(O_1,1), \tag{20}$$

$$\delta_1(i,d)=0,\quad d>1; \tag{21}$$

Recursion: for $j=1,2,\ldots,N$ and $d=1,2,\ldots,t$; $t=1,2,\ldots,T-1$

$$\delta_{t+1}(j,1)=\max_{1\leqslant\tau\leqslant t}\ \max_{1\leqslant i\leqslant N}[\delta_t(i,\tau)a_{ij}]b_j(O_{t+1},1),\quad i\neq j, \tag{22}$$

$$\delta_{t+1}(j,d+1)=\delta_t(j,d)a_{jj}b_j(O_{t+1},d+1), \tag{23}$$

$$\Delta_t(j,i)=\arg\max_{1\leqslant\tau\leqslant t-1}[\delta_{t-1}(i,\tau)a_{ij}],\quad 1\leqslant i\leqslant N,\ i\neq j, \tag{24}$$

$$\Psi_t(j)=\arg\max_{1\leqslant i\leqslant N}[\delta_{t-1}(i,\Delta_t(j,i))a_{ij}],\quad i\neq j, \tag{25}$$

where $\Delta_t(j,i)$ gives the duration in state i just before transiting to state j at time $t-1$ along the best path. Here, $\Psi_t(j)$ gives the previous state index before reaching state j at time t along the best path. These two back pointers are required in the state sequence back tracking. The probability of the best sequence of states is given by

$$P^*=\max_{1\leqslant i\leqslant N}[\delta_T(i,\eta(i))] \tag{26}$$

and

$$\eta(i) = \arg \max_{1 \leq \tau \leq T-1} [\delta_T(i,\tau)], \quad 1 \leq i \leq N, \tag{27}$$

which gives the duration spent in states i at time T along the best path. The final state of the best state sequence is then given by

$$q_T^* = \arg \max_{1 \leq i \leq N} [\delta_T(i, \eta(i))]. \tag{28}$$

The best state sequence is backtracked as follows:
Initialization:

$$x = \eta(q_t^*), \quad t = T, \quad y = x. \tag{29}$$

The following equations are recursed until $t = 1$ is reached:

$$q_{t-x+k}^* = q_t^*, \quad k = 1,2,\ldots,y-1, \tag{30}$$

$$q_{t-x}^* = \psi_{t-x+1}(q_t^*), \tag{31}$$

$$y = \Delta_{t-x+1}(q_t^*, q_{t-x}^*), \tag{32}$$

$$t \leftarrow t - x, \tag{33}$$

$$x \leftarrow y. \tag{34}$$

The array q^* gives the best state sequence.

In our phone recognition experiment using T-HMM, only the Viterbi algorithm is used for maximum likelihood (ML) state sequence decoding. For estimating the parameters of T-HMM based phone recognizer, a different method is employed using the labeled training set, which will be discussed in Sec. IV. In a recent paper, Deng *et al.* (1994) also consider nonstationary HMM states for better phone modeling. They use a polynomial regression function of time to fit the means of the Gaussian observation pdfs at consecutive time intervals in a particular state. The polynomial regression function is supposed to represent the trend of spectral variation of the phone. It may be noted that the mean is only a gross measure of the nature of the pdf and the changing mean model of the nonstationary spectrum would not be a very good approximation of the spectral change. In the discrete pdfs considered here there is no such simplification of the trend of the phone spectra. Also, the consecutive observation pdfs are estimated separately without any polynomial relationship between adjacent pdfs so that any arbitrary spectral change in the phone is well modeled.

II. PHONE DURATION MODELING

In the present phone recognizer since a single state is modeling a phone, the phone duration modeling is equivalent to HMM state duration modeling. In a standard HMM the (implicit) probability of state duration decreases exponentially, which is not suited for most of the phones. Because of this, the phone recognizer would have a bias toward short duration phones than longer duration phones.

There are many solutions for better state duration modeling in HMMs (Ramesh and Wilpon, 1992; Rabiner, 1989; Russel and Moore, 1983; Russel and Cook, 1987; Levinson, 1986). Of these, the inhomogeneous HMM (I-HMM), proposed by Ramesh and Wilpon (1992) is more general and requires less computation compared to the semi-HMM

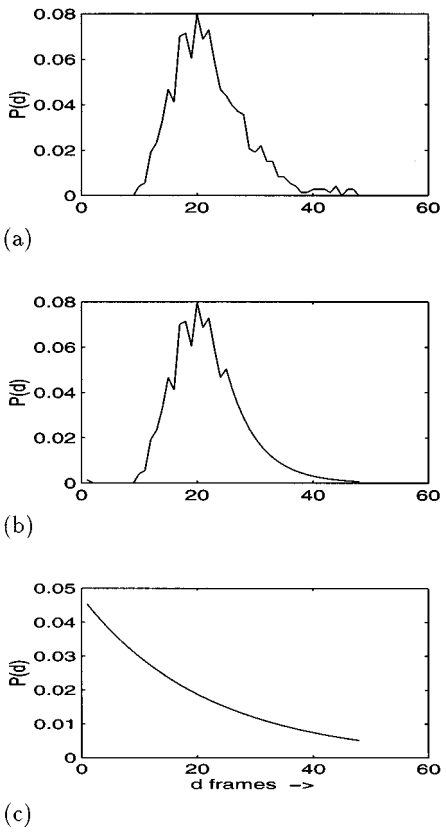


FIG. 3. (a) Actual duration distribution of phone /aw/ calculated from TIMIT. (b) Duration distribution as modeled by an I-HMM state with $D = 25$. (c) Duration distribution as modeled by a standard HMM state.

(Rabiner, 1989). The I-HMM structure is quite similar to the T-HMM, thus enabling the training and testing of both the properties together.

In I-HMM, the state occupancy probability is better modeled because of duration-dependent state transition probabilities: i.e., the transition probability is made dependent on the duration d spent in the state i; i.e., it becomes $a_{ij}(d)$, $1 \leq i, j \leq N$ and $1 \leq d \leq D$ instead of the stationary a_{ij}. Thus each state will have a sequence of transition probability distributions, one for each time instant until a maximum duration limit of D. For duration $d \geq D$, the transition probabilities are truncated to $a_{ij}(d) = a_{ij}(D)$, which provides for an exponential distribution beyond D as in ordinary HMM. The parameter D can be chosen differently for each state based on the speech unit it is modeling. Thus an I-HMM can model state duration with arbitrary probability distribution in the interval $1 \leq d < D$, as given by the following equations:

$$\begin{aligned} P_i(d) &= \left[\prod_{l=1}^{d-1} a_{ii}(l)\right](1 - a_{ii}(d)), \quad 1 \leq d \leq D \\ &= \left[\prod_{l=1}^{D} a_{ii}(l)\right](a_{ii}(D))^{d-D-1}(1 - a_{ii}(D)), \quad d > D. \end{aligned} \tag{35}$$

Figure 3(a) shows the actual duration distribution of the phone /aw/ and Fig. 3(b) and (c) shows the same duration distribution as best could be characterized by an inhomogeneous-HMM state and a standard HMM state, respectively. It is clear that I-HMM can model the duration distribution very effectively. In Fig. 3(b) after $d \geq 25$ frames,

i.e., $D=25$, the distribution becomes geometric as $a_{ij}(d)=a_{ij}(D)$, and constant for all $d \geqslant D$. Since most of the phone durations are approximately exponential after the peak value, D can be fixed slightly greater than the mean duration to model the phone duration effectively. Solutions to all the three standard problems of I-HMM are available in Ramesh and Wilpon (1992).

Now, it is straightforward to include duration-dependent transition probabilities of I-HMM into T-HMM, resulting in TI-HMM. The TI-HMM models both the duration and temporal structure of the phone in a single state, as required. The solutions to the three standard problems of TI-HMM can be easily extended from that of T-HMM given earlier, the main differences being the usage of $a_{ij}(d)$ instead of a_{ij}. The equations which differ from T-HMM to TI-HMM are indicated in the Appendix.

III. COARTICULATION MODELING

It is well known that consonants induce specific formant transitions in the adjacent vowels, which is an important cue for the recognition of consonant itself. The manifestation of some phones like vowels vary drastically because of the neighboring phones, but their manifestation is specific in specific contexts. For phone recognition, this type of contextual dependency is modeled by designing separate LR-HMMs (left-to-right HMM) for each context (Lamel and Gauvain, 1993), resulting in triphone models. Such an approach results in a very large network of HMMs causing very high complexity for sentence decoding. The high complexity is resolved using suboptimal search strategies. In the present approach, since only one state is assigned for each phone, the complexity is limited, but the state also has to capture the contextual effects. This can be achieved by maintaining a different observation distribution for each phonetic context, in a particular state.

Left phone contextual effect can be incorporated into the HMM by making the observation probability distribution of a state j, dependent on the previous state i, from which a transition took place into the present state; i.e., $b_j(k)$ of standard HMM is generalized to $b_{ij}(k)$. Once the transition takes place from state i to state j, the observation distribution used in state j will be $b_{ij}(k)$. This distribution $b_{ij}(k)$ is used as long as the state j is occupied, with the condition that the previous non-self-state is i. Each state will have N context-dependent observation distributions including one nonspeech left context distribution $b_{0j}(k)$, where N equals the number of states (equal to phones) in the model. Note that $b_{jj}(k)$ distribution does not exist. The distribution $b_{0j}(k)$ indicates the probability of observation symbols occurring whenever the phone represented by state j occurs as the beginning phone in the given speech data. This new HMM is referred to as left-contextual-HMM (LC-HMM). The parameter estimation of a phone recognizer using LC-HMM is done similar to that of I-HMM and T-HMM based recognizers, which is explained in the next section. For phone recognition using LC-HMM a fast suboptimal hidden state sequence decoding algorithm is derived below.

A. State sequence decoding for LC-HMM

Given the observation sequence $\{O_1,O_2,O_3,\ldots,O_T\}$ and an LC-HMM λ, let $\delta_t(i)$ be the joint probability of the best state sequence $\{q_1,q_2,\ldots,q_t\}$ with $q_t=i$ and partial observation sequence $O^t=O_1,O_2,\ldots,O_t$. Thus

$$\delta_t(i)=\max_q\{P(q_1,\ldots,q_t=i,O^t|\lambda)\}, \tag{36}$$

where $\lambda=(\mathbf{A},\mathbf{B},\mathbf{\Pi},N,M)$. Here, $\mathbf{A}$, $\mathbf{\Pi}$, M, and N have the same standard meaning as in HMM (Rabiner, 1989). $\mathbf{B}$ denotes the set of context-dependent observation probability distributions, $b_{0j}(k), b_{ij}(k)$ for $1 \leqslant i,j \leqslant N (i \neq j)$ and $1 \leqslant k \leqslant M$. The $\delta_t(i)$ are computed as shown below.

Initialization: for $1 \leqslant i \leqslant N$

$$\delta_1(i)=\pi_i b_{0i}(O_1), \tag{37}$$

$$\Psi_1(i)=0, \tag{38}$$

$$l_1(i)=0. \tag{39}$$

Here, $\Psi_t(i)$ is the back pointer array, which keeps track of the state index in the previous time instant along the best path, ending in state i at time t, and $l_t(i)$ keeps track of the previous non-self-state from where the transition took place into the present state i, in the recent past, from when it stayed in the present state i until time t, along the best path.

Recursion: for $t=1,2,\ldots,T-1$; $j=1,2,\ldots,N$

$$\delta_{t+1}(j)=\max_{1\leqslant i\leqslant N}\Big[\delta_t(i)a_{ij}b_{ij}(O_t), \delta_t(i)a_{ij}b_{l_t(i)j}(O_t)\Big] \quad \begin{matrix} i\neq j \\ i=j \end{matrix} \tag{40}$$

$$\Psi_{t+1}(j)=\arg\max_{1\leqslant i\leqslant N}\Big[\delta_t(i)a_{ij}b_{ij}(O_t), \delta_t(i)a_{ij}b_{l_t(i)j}(O_t)\Big] \quad \begin{matrix} i\neq j \\ i=j \end{matrix} \tag{41}$$

In the above two equations, when $i=j$, i.e., a self-transition occurs, the left context is a non-self-state (phone) from where the transition took place into the present state in the recent past. The left context observation distribution to be used is given by $b_{l_t(j)j}(k)$. The non-self-state back pointer $l_t(i)$ is updated as follows:

$$\text{if } \Psi_{t+1}(j)=j \text{ then } l_{t+1}(j)=l_t(j)$$
$$\text{else } l_{t+1}(j)=\Psi_{t+1}(j). \tag{42}$$

Termination: The probability of the best state sequence is

$$P^*=\max_{1\leqslant i\leqslant N}[\delta_T(i)]. \tag{43}$$

The final best state:

$$q_T^*=\arg\max_{1\leqslant i\leqslant N}[\delta_T(i)]. \tag{44}$$

Backtracking: for $t=T-1,T-2,\ldots,1$

$$q_t^*=\Psi_{t+1}(q_{t+1}^*). \tag{45}$$

The array q^* gives the best state sequence.

The main advantage of introducing context dependency on a phone state is that the variance of the observation symbol pdf will reduce compared to the original context-

independent pdf. This is preferred in a statistical recognition scheme. This idea can be extended to both the left and right context, i.e., triphone models, by making the state observation pdfs depend on both the neighboring states (phones). In the present study, only the left context HMM is evaluated. In any general speech application, for training word models using LC-HMM, Viterbi training can be used. The structure of this context-dependent HMM is similar to the trend-HMM and inhomogeneous-HMM, so that it is possible for all the three phonetic properties to be incorporated together into a state resulting in LCTl-HMM. Each phone may have a different trend in different contexts; LCTl-HMM can capture this context-dependent trend apart from modeling phone durations.

IV. EXPERIMENTS

Phone recognition experiments are conducted using the new HMMs with discrete pdfs. To provide a performance benchmark, phone recognition experiment is also conducted using a standard HMM. As mentioned earlier, the general structure of the phone recognizer using each of the new HMMs consists of a single large ergodic model with the number of states equal to the number of the phones. Each state in the model represents a phone. The parameters of these models are determined from the phonetically labeled speech database, as will be described below. The phone recognition of a given test speech is achieved by the ML state sequence decoding of the ergodic model.

In all the experiments, since the primary motivation is to determine the effectiveness of the new HMMs, no effort is made to achieve best recognition scores by improving the acoustic front-end. All the experiments used only a minimum possible feature set, i.e., weighted cepstral coefficients. Differential features, etc., which can increase the overall phone recognition accuracy of all the models are not attempted. Also, use of continuous observation density for the new HMMs can increase the overall phone recognition performance.

A. Database and speech processing

The speech data used in all the experiments are taken from the phonetically labeled, speaker independent, continuous speech database TIMIT. The training set consisted of 3696 sentences of 462 speakers, taken from all the 8 dialects of the TIMIT training set. The two SA sentences provided for each speaker were not included in the experiments. The core-test set of the TIMIT database containing 192 sentences is used for testing purpose. There are 24 speakers in the core-test set; they are different from those in the training set and they span over all the 8 dialects.

The sampling frequency of the signal is 16 kHz. The speech signal is analyzed in frames of 16 ms with an overlap of 8 ms between frames. Eighteen LPC derived weighted cepstral coefficients are computed every frame, after windowing with a Hamming window and pre-emphasizing with a factor of 0.95. The weighted cepstral feature vectors are quantized using a 256 size codebook, which is designed using the LBG algorithm (Linde *et al.*, 1980).

TABLE I. 60 phone set used in all the experiments.

Phone	Example	Phone	Example	Phone	Example
/b/	bee	/ng/	sing	/ao/	bought
/d/	day	/em/	bottom	/oy/	boy
/g/	gay	/en/	button	/ow/	boat
/p/	pea	/nx/	winner	/uh/	book
/t/	tea	/l/	lay	/uw/	boot
/k/	key	/r/	ray	/ux/	toot
/dx/	dirty	/w/	way	/er/	bird
/q/	bat	/y/	yacht	/ax/	about
/jh/	joke	/hh/	hay	/ix/	debit
/ch/	choke	/hv/	ahead	/axr/	butter
/s/	sea	/el/	bottle	/ax-h/	suspect
/sh/	she	/iy/	beet	/pau/	pause
/z/	zone	/ih/	bit	/epi/	silence
/zh/	azure	/eh/	bet	/h#/	non-speech
/f/	fin	/ey/	bait	/bcl/	*b* closure
/th/	thin	/ae/	bat	/dcl/	*d* closure
/v/	van	/aa/	bott	/gcl/	*g* closure
/dh/	then	/aw/	bout	/pcl/	*p* closure
/m/	mom	/ay/	bite	/tcl/	*t* closure
/n/	noon	/ah/	but	/kcl/	*k* closure

B. Parameter estimation

All the parameters of the phone recognizer are determined using the statistical information of the corresponding events in the speech database. The number of states N in the ergodic model (using either HMM, I-HMM, T-HMM, TI-HMM, or LC-HMM) is fixed as 60 in all the experiments, with one state representing a phone. The 60 phones used here are taken from the 61 phone set of TIMIT (Table I); the phones /ng/ and /eng/ are merged into a single phone.

The parameters of the phone recognizer using the standard HMM $\lambda=(\mathbf{A},\mathbf{B},\mathbf{\Pi})$ are calculated as follows: As the speech database is phonetically labeled, the labeling information is first mapped onto the codeword sequences of the entire speech database. With this it is straight forward to construct histograms of codewords occurring in each phone and by normalizing, the observation probability distributions of all the states of HMM (i.e., all phones) are found. By counting the number of times each phone occurs in the beginning of all sentences in the training set, the initial state occupancy probability distribution is determined. Similarly, by counting the number of times a phone (frame) transits to another phone (frame), including itself, the transition probability distribution of the state characterizing that phone is found, and this is repeated for all states.

The design of phone recognizers using T-HMM, I-HMM, TI-HMM, and LC-HMM is similar to the design given above. Determining the parameters of T-HMM is similar to that of HMM parameters, except for the duration-dependent observation probabilities. The $b_j(k,d)$ are determined by counting the codewords at different time instants in all occurrences of the phones, as described below.

Let N_d be the total number of occurrences of any codeword in all the occurrences of the phone at the dth frame and let N_d^k be the number of occurrences of the codeword k at the dth frame in all N_d occurrences of the phone. Thus $b_j(k,d)=N_d^k/N_d$, where j is the phone index. The $b_j(O_t,d)$ used in Sec. I A corresponds to $b_j(k,d)$ for O_t

$=v_k$. The probabilities are evaluated for all codewords $1 \geqslant k \geqslant M$ and for all durations $1 \geqslant d \geqslant D-1$. The last observation distribution $b_j(k,D)$, which represents the observation symbol probabilities for all durations $d \geqslant D$ of the phone is designed according to the equation:

$$b_j(k,D) = \sum_{d=D}^{\text{End}} N_d^k \Big/ \sum_{d=D}^{\text{End}} N_d . \tag{46}$$

Here the label End denotes the ending frame of the phone tokens used. A fixed value of $D=10$ for all the phones is found to give satisfactory phone recognition performance.

In the I-HMM based phone recognizer, all parameters are determined similarly to HMM, except $a_{ij}(d)$ which are similar to the $b_j(k,d)$ in T-HMM. The duration-dependent transition probabilities $a_{ij}(d)$ are determined by counting the occurrences of phone transitions from phones represented by state i to phones represented by state j, with the condition that the transition took place when the first phone duration is exactly d frames. Let $M_d(i,j)$ be the number of occurrences of phone pairs represented by state i and state j in that order, in which the first phone i has occurred for a duration d. Let $M_d(i)$ be the number of occurrences of phone i for duration $\geqslant d$ frames; then

$$a_{ij}(d) = M_d(i,j)/M_d(i), \quad \text{for } 1 \leqslant i,j \leqslant N; \ 1 \leqslant d \leqslant D-1. \tag{47}$$

The last transition probability distribution $a_{ij}(D)$ which represents the transition probabilities for all durations $d \geqslant D$ is estimated as given below. Here all the transitions, where the first phone duration is $\geqslant D$, are included in determining the $a_{ij}(D)$, i.e.,

$$a_{ij}(D) = \frac{\sum_{d=D}^{\text{End}} M_d(i,j)}{M_D(i)}. \tag{48}$$

These equations are repeated for all the states. The value of $D=25$ is found to give optimum results in our experiment. The parameter estimation of the TI-HMM based recognizer directly follows from that of T-HMM and I-HMM, respectively.

The phone recognizer using LC-HMM also has only one parameter set different from standard HMM, i.e., $b_{ij}(k)$. These left phone context-dependent observation distributions are determined by counting the occurrences of all the codewords in the phone occurrences when they have occurred with the specific left context. Let $N_{ij}(k)$ be the number of occurrences of codeword k, in all the occurrences of phone j with the left context as phone i. And let N_{ij} be the number of occurrences of any codeword in phone j with the left context as phone i. Hence, $b_{ij}(k) = N_{ij}(k)/N_{ij}$. This is repeated for all codewords and phones with all possible left contexts, including the nonspeech context.

C. Parameter smoothing

In all the new models, the number of parameters to be estimated has increased compared to a standard HMM. For a reliable estimate of these parameters, the events corresponding to each parameter have to occur in a large number in the speech database, which may not be possible. To improve the estimate of such rarely occurring events, various smoothing techniques have been explored. In the standard HMM based phone recognizer, all the $b_j(k)$ are restricted to a minimum value of $\epsilon = 0.001$, which has been found to be optimum. However in T-HMM, since the observation distribution has become duration dependent, the sample histograms derived have more zeros. Instead of applying a fixed floor (ϵ) to the distribution, the pdfs at different durations can be smoothed. This will compensate for the lack of sufficient training data. Representing the smoothed distribution as $\overline{b_j}(k,d)$,

$$\overline{b_j}(k,d) = \sum_{l=0}^{2} h(l) b_j(k,d+l), \quad \forall d,k. \tag{49}$$

The above three-point moving average filter with $h(l) = \frac{1}{3}$; $\forall l$ is found to be satisfactory.

Similarly the duration-dependent transition probabilities $a_{ij}(d)$ in I-HMM requires many occurrences of possible phone pairs, for all durations of the first phone. But some durations may not have enough occurrences; they are interpolated from neighboring $a_{ij}(d)$'s, as in Eq. (50) with filter coefficients $h(l) = \frac{1}{3}$, $\forall l$, i.e.,

$$\overline{a_{ij}}(d) = \sum_{l=0}^{2} h(l) a_{ij}(d+l), \quad \forall d. \tag{50}$$

In LC-HMM, each phone occurrence with a particular left context may be infrequent, resulting in poor estimate of the left context-dependent observation probabilities. One way to counter this problem is to first compute a context-independent pdf, $b_j(k)$ of the phone, which has a larger number of events, and hence a better estimate of the pdf. The context-dependent pdfs are then estimated as a weighted combination of the context-independent pdf, as shown below, i.e.,

$$\overline{b_{ij}}(k) = \zeta b_{ij}(k) + (1-\zeta) b_j(k). \tag{51}$$

Here the weighting factor ζ is a function of the total number of codewords, K_{ij} occurring in the phone j for the context i, i.e., $\zeta = K_{ij}/K_j$, where $K_j = \Sigma_{i=1}^{N} K_{ij}$.

D. Recognition results

The phone recognition of the test set sentences using each of the new models is obtained using their respective Viterbi algorithms. The decoded phone string is compared with the original labeled phone string of the database using dynamic programming and the performance scores are computed by obtaining the number of phones correctly recognized, inserted, deleted, and substituted. While reporting results in Table II, since some phone confusions are not

TABLE II. Phone recognition results on core-test set.

Model used after smoothing	% correct (40 classes)
HMM	44.48%
T-HMM	46.42%
I-HMM	47.58%
LC-HMM	46.72%
TI-HMM	49.91%

serious, a reduced phone set of 40 phone classes is chosen, which includes 39 classes as in Lee and Hon (1989), plus the glottal stop /q/, which is treated as a separate class. The phone groups in which the confusions are allowed are {bcl, dcl, gcl, pcl, tcl, kcl, epi, pau, h#}, {sh, zh}, {m, em}, {n, en, nx}, {l, el}, {hh, hv}, {ih, ix}, {aa, ao}, {ah, ax, ax-h}, {uw, ux}, and {er, axr}.

From Table II it is clear that there is improvement in the percentage of phones correctly recognized by the addition of each degree of freedom to the HMM. This shows the importance of each of the phonetic property in speech recognition and, second, that they can be incorporated as additional parameters in the HMM. Recognition accuracy of almost every phone has shown improvement in the T-HMM compared to HMM, with diphthongs {ay, ey, aw} accounting for maximum improvement, in the range of 10%–12%. With incorporation of duration (I-HMM, TI-HMM), long vowels {aa, iy, aw}, and short vowels {ih, ah, uh, er} have shown remarkable improvement, followed by other phones such as affricates, fricatives, semi-vowels, nasals, and stops. The left context modeling (LC-HMM) helped phones such as /er/, /ih/, /iy/, /l/, /m/, /n/, etc., which are otherwise susceptible. By excluding the glottal stop /q/ from the database as done in Lee and Hon (1989), the absolute accuracy can be increased by 2%–3%, as they were almost never recognized in any experiment. In all the cases the number of phones inserted (insertion error) was maintained less than 9%.

If the D in T-HMM is chosen smaller than the maximum duration of the phone being modeled by that state, the temporal structure of the phone up to a duration D is captured well, but beyond D only one observation pdf $b_j(k,D)$ models the whole phone segment. The choice of D should depend on the maximum duration of phone but it has to be chosen considering the availability of enough training data.

E. Discussion

The lacunae of a standard HMM in modeling the duration and context (temporal dependency), etc., has been well recognized. There are several approaches in the literature to tackle these problems. One approach, such as in SPHINX (Lee *et al.,* 1990), incorporates the phonetic properties by using different HMMs for each triphonic context and applying duration penalty while decoding, resulting in a very large and complex HMM network. In another approach (Ljolje and Levinson, 1991; Ljolje, 1994) the HMM state is made more powerful by adding extra degrees of freedom to incorporate the phonetic properties. The difference between the two is, in the first approach, knowledge is added to the recognition system in the form of a complex network, but the basic HMM form is retained to be the same; whereas, in the second approach, the basic HMM is made more powerful so that the overall recognition system will be a simple ergodic network, and the statistical parameters of the resulting model can be optimized using all the phonetic properties simultaneously. In the first approach the parameters of the duration and acoustic models are optimized independently. The present paper falls into the second category, where the emphasis is on making a single HMM state to characterize a phone. In a series of papers Ljolje *et al.* (1991, 1994) worked in this direction; initially they used a CVD-HMM (Levinson, 1986) where an extra pdf in each state models the phone/state duration, dispensing with state's self-loop. The duration-dependent transition probabilities of I-HMM and TI-HMM used in our present work help toward better phonetic decoding as they also carry information of contextual effects on the duration of phones. For example, a contention whether a phone j or a phone k should follow phone i, can occur in a CVD-HMM based phone decoding if transition (bigram) probabilities a_{ij} and a_{ik} are comparable; but if phone i has different duration pdfs in the context j and k, respectively, then $a_{ij}(d)$ and $a_{ik}(d)$ of I-HMMs will resolve this by proving which phone is more probable when the initial phone i has occurred for d time instants. Note that $\Sigma_{j=1}^{N} a_{ij}(d)=1$ (j varying across all contexts), for any d in I-HMM. This is a novel way of capturing durational information of a phone in a context, rather than maintaining separate duration pdfs for each context.

Although a CVD-HMM state characterizes the duration of a phone, its single observation pdf is not sufficient to capture the spectral variations of a phone, as discussed earlier. Recognizing this, in his later paper Ljolje (1994) uses three sequential distributions (similar to a LR-HMM) in the CVD-HMM state to characterize the spectral variations of a phone. Ljolje (1994) models each of the properties (acoustics, phonotactics, duration, etc.) independently and also gives extra weight to some of the properties while decoding. Whereas in our work, for catering multiple observation pdfs for each phone, we developed new models such as TI-HMM, with complete mathematical solutions which also capture the duration implicitly. And, we use the actual probabilities of each phonetic property without fixing any arbitrary weightage.

Ljolje (1994) proposed a novel way of context modeling called quasitriphonic model with good success. We have shown how to incorporate the left contextual effect in a HMM (LC-HMM), retaining the single state-to-phone relationship, thus maintaining the simplicity of the original system. This leaves scope for extending the context dependency to triphonic level, utilizing ideas like quasitriphonic models, generalized triphones (Lee *et al.,* 1990), etc., yet maintaining the simplicity of a single state-to-phone relationship.

In our new HMMs, the number of model parameters has significantly increased and some novel methods to effectively estimate the parameters from the given amount of training data have to be developed (Sitaram and Sreenivas, 1997); this work is presently being carried out with good success. As this approach consists of a single ergodic model, many other phonetic properties, such as trigram (Lamel and Gauvain, 1993; Ljolje, 1994), can also be easily incorporated into our system.

V. CONCLUSIONS

Recognizing the weakness of a standard HMM for phone recognition, which is crucial for the success of continuous speech recognition, in this paper we incorporate three important properties of a phone into the HMM. We have proposed three new HMMs, in each case making the

HMM state more powerful. Algorithms for parameter estimation and Viterbi decoding of the new HMMs are presented. Experiments with the TIMIT database show the importance of phonetic properties for the recognition accuracy. This study opens a new line of thought to find useful properties of speech and to try to incorporate them in HMMs. By using a bigger training set and newer parameter estimation techniques, we feel the proposed models will perform better.

APPENDIX: SOLUTIONS OF TI-HMM

The algorithms of T-HMM given in Sec. I A are applicable to TI-HMM with the following changes: In the forward–backward algorithm Eqs. (4) and (5) become

$$\alpha_{t+1}(j,1)=\sum_{\tau=1}^{t}\sum_{i=1}^{N}\alpha_t(i,\tau)a_{ij}(\tau)b_j(O_{t+1},1),\quad i\neq j, \tag{A1}$$

$$\alpha_{t+1}(j,d+1)=\alpha_t(j,d)a_{jj}(d)b_j(O_{t+1},d+1), \tag{A2}$$

and Eq. (9) becomes

$$\beta_t(j,d)=\left[\sum_{i=1,i\neq j}^{N}\beta_{t+1}(i,1)a_{ji}(d)b_i(O_{t+1},1)\right]+\beta_{t+1}(j,d+1)a_{jj}(d)b_j(O_{t+1},d+1). \tag{A3}$$

In the re-estimation formulae Eqs. (15) and (16) become

$$\xi_{t\ i\neq j}(i,j,d)=\frac{\alpha_t(i,d)a_{ij}(d)b_j(O_{t+1},1)\beta_{t+1}(j,1)}{P(O/\lambda)}, \tag{A4}$$

$$\xi_{t\ i=j}(i,i,d)=\frac{\alpha_t(i,d)a_{ii}(d)b_i(O_{t+1},d+1)\beta_{t+1}(i,d+1)}{P(O/\lambda)}. \tag{A5}$$

Unlike Eq. (17) the transition probabilities of TI-HMM are re-estimated as follows:

$$\text{for } d<D\quad a_{ij}(d)=\sum_{t=1}^{T-1}\xi_t(i,j,d)\Big/\sum_{t=1}^{T-1}\gamma_t(i,d),\quad 1\leqslant i,j\leqslant N; \tag{A6}$$

$$a_{ij}(D)=\sum_{t=1}^{T-1}\sum_{d=D}^{T}\xi_t(i,j,d)\Big/\sum_{t=1}^{T-1}\sum_{d=D}^{T}\gamma_t(i,d),\quad 1\leqslant i,j\leqslant N. \tag{A7}$$

And in Viterbi algorithm, Eqs. (22)–(25) become:

$$\delta_{t+1}(j,1)=\max_{1\leqslant\tau\leqslant t}\ \max_{1\leqslant i\leqslant N}[\delta_t(i,\tau)a_{ij}(\tau)]b_j(O_{t+1},1),\quad i\neq j, \tag{A8}$$

$$\delta_{t+1}(j,d+1)=\delta_t(j,d)a_{jj}(d)b_j(O_{t+1},d+1), \tag{A9}$$

$$\Delta_t(j,i)=\arg\max_{1\leqslant\tau\leqslant t-1}[\delta_{t-1}(i,\tau)a_{ij}(\tau)],\quad 1\leqslant i\leqslant N,\ i\neq j, \tag{A10}$$

$$\Psi_t(j)=\arg\max_{1\leqslant i\leqslant N}[\delta_{t-1}(i,\Delta_t(j,i))a_{ij}(\Delta_t(j,i))],\quad i\neq j. \tag{A11}$$

Crystal, T. H., and House, A. (**1982**). ''Segmental durations in connected speech signals: Preliminary results,'' J. Acoust. Soc. Am. **72,** 705–716.

Deng, L., Aksmanovic, M., Sun, X., and Wu, C. J. J. (**1994**). ''Speech recognition using hidden Markov models using polynomial regression functions as nonstationary states,'' IEEE Trans. Speech Audio Process. **2,** 507–520.

Lamel, L. F., and Gauvain, J. L. (**1993**). ''High performance speaker independent phone recognition using CDHMM,'' Proceedings of the 3rd European Conference on Speech, Communications, and Technology, Vol. 1, pp. 121–124.

Lee, K. F., and Hon, H. (**1989**). ''Speaker-independent phone recognition using hidden Markov models,'' IEEE Trans. Acoust. Speech Signal Process. **37,** 1641–1648.

Lee, K. F., Hon, H., and Reddy, R. (**1990**). ''An overview of the Sphinx speech recognition system,'' IEEE Trans. Acoust. Speech Signal Process. **38,** 35–45.

Levinson, S. E. (**1986**). ''Continuously variable duration hidden Markov models for automatic speech recognition,'' Comput. Speech Lang. **1,** 29–46.

Linde, Y., Buzo, A., and Gray, R. M. (**1980**). ''An algorithm for vector quantizer design,'' IEEE Trans. Commun. **COM-28**, 84–95.

Ljolje, A. (**1994**). ''High accuracy phone recognition using context clustering and quasi-triphonic models,'' Comput. Speech Lang. **8,** 129–151.

Ljolje, A., and Levinson, S. E. (**1991**). ''Development of an acoustic-phonetic hidden Markov model for continuous speech recognition,'' IEEE Trans. Signal Process. **39,** 29–39.

Rabiner, L. R. (**1989**). ''A tutorial on hidden Markov models and selected applications in speech recognition,'' Proc. IEEE **77,** 257–285.

Ramesh, P., and Wilpon, J. G. (**1992**). ''Modeling state durations in hidden Markov models for automatic speech recognition,'' Proc. IEEE Int. Conf. Acoust. Speech Signal Process. **1**, I-381–I-384.

Robinson, T., and Fallside, F. (**1991**). ''A recurrent error propagation network speech recognition system,'' Comput. Speech Lang. **5,** 259–274.

Russel, M. J., and Cook, A. (**1987**). ''Experimental evaluation of duration modeling techniques for automatic speech recognition,'' Proc. IEEE Int. Conf. Acoust. Speech Signal Process. **4**, 2376–2379.

Russel, M. J., and Moore, R. K. (**1985**). ''Explicit modelling of state occupancy in hidden Markov models for automatic speech recognition,'' Proc. IEEE Int. Conf. Acoust. Speech Signal Process. **1**, 5–8.

Sitaram, R. N. V., and Sreenivas, T. V. (**1997**). ''Scalar quantization in discrete hidden Markov models for better performance,'' Int. Conf. Speech Process., Seoul, Korea, 26–28 Aug. 1997 (accepted for presentation).

Simple model of a piano soundboard

N. Giordano
Department of Physics, 1396 Physics Building, Purdue University, West Lafayette, Indiana 47907-1396

(Received 18 September 1996; revised 3 April 1997; accepted 8 April 1997)

The vibrational properties of a simple finite-element model of a piano soundboard are considered. The main focus is on the behavior of the mechanical impedance in the musically important frequency range $\sim 50-10^4$ Hz. The model includes the effects of elastic anisotropy and the ribs. It is argued that the ribs are an essential ingredient for producing the behavior of the impedance which is observed experimentally. © *1997 Acoustical Society of America.* [S0001-4966(97)01008-4]

PACS numbers: 43.75.Mn [WJS]

INTRODUCTION

The problem of modeling a piano would at first glance seem to be straightforward. One needs to consider the vibrations of a flexible string which is held fixed at one end and is attached to a flexible plate, the soundboard, at the other. The vibrations of the soundboard lead in turn to sound production. However, while the physics of vibrating strings and plates is generally well understood, there are numerous complications, such as the effects of string stiffness, the precise nature of the connection of the string to the soundboard (which leads to coupling of the different modes of polarization of the vibrations of the string), the nonlinear behavior of the hammers, and the complicated shape of the soundboard. Because of such difficulties, the current state of the art in piano modeling does not produce a convincing piano sound; nevertheless, progress has been made toward this goal in recent years. For example, Chaigne and Askenfelt[1,2] have treated the hammer–string interaction in detail, along with the effect of string stiffness, and have obtained good agreement with certain aspects of the experimental observations. However, their calculations do not include coupling to the soundboard, assuming instead that both ends of the string are held perfectly rigid. Energy loss from the string is then accounted for through the use of *ad-hoc* frequency-dependent damping terms which are added to the equation of motion for the string. This approach has serious limitations, as it precludes, for example, any treatment of coupling between strings.

One would clearly like to model the soundboard in a quantitative manner, and while progress along these lines has been made for the piano,[3–7] the guitar,[8–11] and the violin,[12] much remains to be done in this area. The purpose of the present paper is to discuss a simple model of a piano soundboard, and compare its properties with those observed for real soundboards. The model that we ultimately arrive at will be an extremely simple one consisting of a square board, with the grain oriented parallel to one edge of the board and with ribs running perpendicular to the direction of the grain. This is certainly not intended to be a ''completely realistic'' model of a real soundboard, but rather as an aid in understanding those aspects of the board which are essential for its proper acoustical performance. One of the goals of this work is to construct a real-time model (as opposed to one which operates in the frequency domain) which can be used in conjunction with piano string simulations like those of Ref. 1. We will also attempt to identify how different properties of the model, such as elastic anisotropy and the presence of ribs, affect the behavior. Our calculations and arguments are not profound, but we believe that they bring out some important points that we have not seen discussed previously in the literature.

I. EXPERIMENTAL FACTS AND SOME GENERAL OBSERVATIONS

Figure 1 shows, in schematic form, results for the magnitude of the mechanical impedance, $|Z| \equiv F_0/v_0$, as a function of frequency from experiments with an upright piano.[13] Here F_0 is the amplitude of an applied force, which is a sinusoidal function of time with frequency f. This force is applied to a particular point on the board (as will be specified below), and v_0 is the steady-state amplitude of the velocity of the board at the same location, also at frequency f. While our sketch shows the impedance as a fairly smooth function of frequency, the effects of soundboard resonances are clearly seen in the measured data, and several such resonances are shown semiquantitatively in Fig. 1. We also show the measured sound power level as a function of frequency, as recorded with a microphone a short distance from the soundboard. Here the board was driven harmonically, and F_0 was adjusted to obtain the same value of v_0 at all frequencies. The bottom plot in Fig. 1 shows the phase of the steady-state soundboard velocity relative to the drive force. We will be comparing all of the results in Fig. 1 with our model calculations below. The results for the impedance in Fig. 1 are the main experimental facts that we wish to understand and explain. It is worthwhile at this point to consider precisely which features of the experimental results are musically important. Let us discuss the frequency ranges <100, $100-10^3$, and 10^3-10^4 Hz separately.

At frequencies which are below the fundamental soundboard mode, $|Z|$ increases as f is lowered. For the particular soundboard considered in Ref. 13 this mode occurred near 100 Hz, and that seems to be a typical value.[3,5,15] The sound generation efficiency in this range will generally be very low, since the size of the source differs greatly from the wavelength of the radiation; this low efficiency is also evident from the measurements (Fig. 1). In addition, the fundamental frequency for most strings is above this range, so they

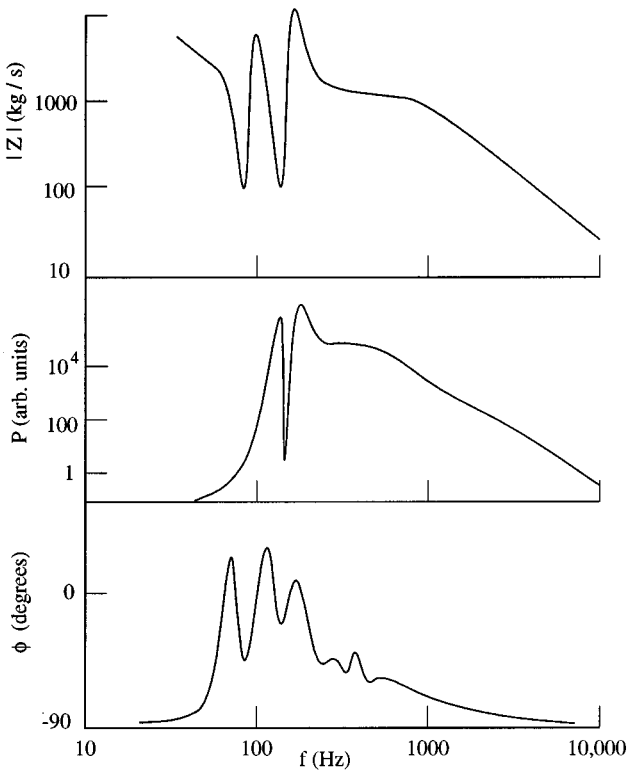


FIG. 1. Schematic soundboard properties after the results given in Ref. 13 for an upright piano. Top: magnitude of the mechanical impedance; middle: sound power level;[14] bottom: phase of the soundboard velocity relative to the phase of the driving force. The resonant structure due to a few of the lowest modes is shown. The results were obtained by exciting the board at a point on the treble bridge near the center of the soundboard, and the velocity was measured at the driving point (the results for other excitation points were similar). Note that the result for the sound level was obtained with a constant velocity at the driving point.

do not drive the soundboard at these frequencies. The behavior of $|Z|$ in this range should also be relatively easy to model, since as we will see below it depends mainly on having the fundamental soundboard mode at the proper frequency.

In the range $100-10^3$ Hz one finds that $|Z|$ exhibits a number of resonances, and it varies around an ''average'' value which is near 10^3 kg/s. This average value depends somewhat (by a factor of ~ 2) on the position at which one measures the impedance (i.e., where the force is applied). The resonant structure in $|Z|$ arises from the many soundboard modes which occur above the fundamental. As is well appreciated,[13,16,17] it is important that this resonant structure not be too pronounced, as it could lead to an uneven response as one moves across the keyboard. The difference between the peaks and valleys in $|Z|$, and the widths of these resonances, will depend on the magnitude of the loss. We should also note that the resonant structure in $|Z|$ varies with location on the soundboard.[13] It seems likely that good pianos are designed to avoid positioning a string at locations where it can strongly excite a resonance corresponding to its particular frequency. A second important feature of the impedance in this frequency range is its overall value. This is important with regards to the strength with which a string couples to the soundboard. The efficiency with which energy is transmitted from the string to the soundboard is governed by the impedance mismatch between the two. This is of great importance, since it determines the rate at which a string vibration decays and hence the duration of the sound. It seems likely that the value of $|Z|$ seen in Fig. 1, which has also been reported elsewhere,[18] gives the ''proper'' amount of coupling and hence the proper decay rate. Here the term ''proper'' means that the coupling and decay rate yield sounds which are musically pleasing to a listener. Thus it is also the behavior observed in concert-quality pianos. A successful soundboard model should therefore give an impedance near 10^3 kg/s in this middle frequency range. This conclusion has also been reached by Conklin.[19]

Lastly we consider the high-frequency range 10^3-10^4 Hz. Here the resonant structure is much less pronounced than at lower frequencies, and there is a monotonic decrease in $|Z|$ from a value near 10^3 kg/s at $\sim 10^3$ Hz to ~ 30 kg/s at 10^4 Hz.[20] We believe that this falloff is important for the following reason. The string impedance varies somewhat across the keyboard, from ~ 10 kg/s in the bass to ~ 5 kg/s in the treble,[13] and is always much less than the impedance of the board. We have already noted that the string-soundboard mismatch determines how fast a string vibration will decay. It is well-known that this decay time becomes shorter as one moves from strings in the bass to those in the treble, and also that for a given string the decay time becomes shorter as one goes to the higher frequencies (higher-order partials).[21,22] Theoretically we know that the decay time will become shorter as the impedance mismatch is reduced. Hence one way to account for this behavior of the decay time is to assume that Z of the soundboard decreases at high frequencies so as to reduce the mismatch with the string. The argument then is that the decrease of the impedance of the soundboard observed at high frequencies is important for obtaining the proper frequency dependence of the decay rate. Given that this is also observed in the experiments of Ref. 13, the implication is that a successful soundboard model must exhibit this behavior. (However, there may be other sides to this argument; we will return to this point below.[23])

Let us now consider what is perhaps the simplest possible model of a soundboard, and compare its properties to those we desire. The board is modeled as a single mass connected to a rigid support via a spring, with a damping term to account for sound generation and losses inherent in the board itself. A plausible equation of motion is then

$$m\frac{d^2z}{dt^2} = -Kz - R\frac{dz}{dt} + F, \tag{1}$$

where z is the displacement (not to be confused with Z), m the mass of the soundboard, K is the spring constant, F is an applied force due to the string, and R determines the magnitude of the damping. The properties of this simple ''single oscillator'' model of the board can be readily calculated. If we assume a driving force of the form $F=F_0e^{i\omega t}$, the steady-state velocity of the board is given by[16]

$$v = \frac{F_0 e^{i\omega t}}{R + i(\omega m - K/\omega)}. \tag{2}$$

The mechanical impedance is $Z \equiv F/v$. From (2) we find that above the resonant frequency of the board ($\omega_0 = \sqrt{K/m}$), the imaginary part of Z increases with frequency, while its real part is constant. Since the energy loss rate is equal to $\mathrm{Re}(Fv) = \mathrm{Re}(F^2/Z)$ [here Re( ) denotes the real part of an expression], one sees that the loss will decrease rapidly at frequencies above ω_0. This high-frequency response is ''mass dominated,'' and it is seen to be property of any mass-spring system at frequencies above all of the resonant frequencies of the system.

For a real piano soundboard the fundamental resonance is typically near 100 Hz, and the resonant frequency of the model (1) can always be adjusted to give this value for ω_0. The other parameters in the model can then be chosen to give an overall value of $|Z|$ near the desired 10^3 kg/s. However, the single oscillator model exhibits an impedance which increases above its resonant frequency, so it will *always* yield a string decay rate that decreases with increasing frequency. If a real soundboard were to behave in this fashion, only the high frequency partials would remain after a short time, yielding an unpleasant tone. This is, of course, contrary to observations.[21,22,16]

The difficulty with the single oscillator model (1) can be restated as follows. The fundamental (i.e., lowest frequency) mode of the soundboard must be near 100 Hz, so as to match the fundamental mode of real soundboards. However, the behavior at frequencies up to at least 10^4 Hz must not be ''masslike.'' The only way this can happen is for the soundboard to have a distribution of resonant modes over this range. It must therefore be treated as a spatially extended system, which should not come as a surprise. A model with only one or a small number of oscillators must therefore be inadequate for a treatment of the behavior at frequencies above a few hundred Hz. This leads us to consider a spatially extended model in the next section.

We should note that a model which employs (1), or something like it, to describe each of a small number of oscillators may be useful in dealing with the low-frequency behavior. Models of this kind for a guitar soundboard have been used successfully in the appropriate low-frequency range.[9–11]

II. ISOTROPIC BOARD

There have been several previous discussions of piano soundboards, from both experimental and theoretical points of view.[24] Among the experiments have been studies of the detailed vibrational modes, the eigenfrequencies and eigenvectors, exhibited by specific soundboards. This work has usually focused on the lowest few modes, typically below a few hundred Hz,[4,5,13,15,19,25] since at higher frequencies the relative separation of the modes becomes small and their displacement patterns (eigenvectors) become more complicated, making the modes difficult to distinguish individually. Theoretically it has been shown that these low-frequency vibrations can be understood quantitatively in terms of models based on vibrating plates, provided that one includes the ribs and bridges. One interesting observation[5] is that the vibrations of the rim must also be included in order to accurately account for the modal frequencies and eigenvectors.

Since our ultimate goal is to produce a soundboard model which can be used for time-domain simulations of the entire string-soundboard-sound system, it is necessary to push the modeling to higher frequencies. We will therefore attempt to focus instead on the general features of the model which are needed to obtain the behavior which was discussed in the previous section. Our aim is not to construct a model which accurately accounts in detail for the vibrational modes of any real soundboard. Rather, we want to understand what ingredients are necessary in order for a model to exhibit the general behavior observed for the impedance as a function of frequency as sketched in Fig. 1. Of course, at some point in the future one may want to design a model which does both.

In this and the next two sections we will consider several models of vibrating plates. The first and simplest model we consider is a homogeneous, isotropic plate. The equation of motion is then[16,26]

$$\rho h \frac{\partial^2 z}{\partial t^2} = -D\nabla^4 z, \tag{3}$$

where z is the displacement of the plate in the direction perpendicular to its plane, h its thickness, ρ its density, and D is the rigidity which is given by

$$D = \frac{h^3 E}{12(1-\nu^2)}. \tag{4}$$

Here, E is Young's modulus, and ν is Poisson's ratio. We consider a thin plate with vibrations only along z, so $\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2$. This model assumes that the plate is elastically isotropic, while a real soundboard certainly is not. We will add anisotropy to the model in the next section, but it is useful to consider the isotropic case first.

While the equation of motion (3) can be solved exactly (for simple plate geometries), we will soon be considering more complicated models for which such solutions are not readily available.[27,28] We will therefore consider the finite difference version of (3) in which the plate is divided into N^2 elements, each of area $(\Delta x)^2$ and mass $\rho h(\Delta x)^2$, arranged on a square grid. The displacement of the plate as a function of position and time then becomes $z(x,y,t) \rightarrow z(i,j,n)$, where i, j, and n are integers with $x = i\Delta x$, $y = j\Delta x$ (for simplicity we take the grid size to be the same in the x and y directions), and $t = n\Delta t$. We then write the derivatives in (3) in the usual finite difference forms,[29] and rearrange to obtain $z(i,j,n+1)$, the displacement at time step $n+1$, in terms of the displacement at earlier times. The result is

$$z(i,j,n+1)=a_1z(i,j,n)+a_2z(i,j,n-1)+a_3[z(i+2,\ j,n)-4z(i+1,\ j,n)+6z(i,j,n)-4z(i-1,\ j,n)+z(i-2,\ j,n)]$$
$$+a_4[z(i,j+2,\ n)-4z(i,j+1,\ n)+6z(i,j,n)-4z(i,j-1,\ n)+z(i,j-2,\ n)]+a_5[z(i+1,\ j+1,\ n)$$
$$+z(i+1,\ j-1,\ n)+z(i-1,\ j+1,\ n)+z(i-1,\ j-1,\ n)-2z(i+1,\ j,n)-2z(i-1,\ j,n)-2z(i,j+1,\ n)$$
$$-2z(i,j-1,n)+4z(i,j,n)]+a_6F(i,j,n), \tag{5}$$

where we have grouped the terms according to their origin (this grouping will also be convenient when we discuss generalizations of the model below). The coefficients a_i are given by

$$a_1=\frac{2}{1+\beta},\quad a_2=\frac{-1+\beta}{1+\beta},$$
$$a_3=\frac{-D(\Delta t)^2}{\rho h(\Delta x)^4},\quad a_4=\frac{-D(\Delta t)^2}{\rho h(\Delta x)^4}, \tag{6}$$
$$a_5=\frac{-2D(\Delta t)^2}{\rho h(\Delta x)^4},\quad a_6=\frac{(\Delta t)^2}{\rho h(\Delta x)^2},$$

with

$$\beta=\frac{R\Delta t}{2\rho h}. \tag{7}$$

The parameter β is a measure of the dissipation, as we have now generalized the model (3) slightly by adding both a driving force and damping. The force applied to element (i,j) at time step n is $F(i,j,n)$, and is allowed to be a function of position. The damping force on a mass element is assumed to be of the form $-Rv(i,j)(\Delta x)^2$, where $v(i,j)$ is its velocity and R is a parameter which sets the scale for the damping [compare with (1)].

For the calculations we began with $z(i,j,0)=0$ for all i and j; i.e., an undisplaced board at time $t=0$. A force $F_0\sin(2\pi ft)$ was then applied at only one grid site so as to have the same situation as in measurements of the driving point impedance. The displacement of the board as a function of time was then computed iteratively using (5). Here and for all of the calculations described below, clamped boundary conditions were employed at the edges of the board so as to approximate a soundboard clamped to a rim. Our main concern will be with the behavior at frequencies well above those of the lowest modes, so we expect that similar results would be found with simply supported or hinged boundaries (the modes at the lowest frequencies are generally the ones affected most by the boundary conditions). At each frequency the simulation was run for several (usually ten or more) periods of the drive frequency to allow the system to reach a steady state, before calculating the amplitude of the velocity at the driving point (used to calculate $|Z|$), the phase of this velocity relative to F, and other quantities.[30]

Some typical results for an isotropic board are given in Fig. 2 where we show the driving point impedance as a function of frequency for two choices of the Young's modulus. Here we consider only $|Z|$; phase information is also available, but we will not consider it until we have arrived at a model which yields acceptable results for $|Z|$ (which will not come in this section). $|Z|$ is the ratio of the amplitude of the force at a particular location to the amplitude of the velocity of the board at the same location. Note that we chose a driving point somewhat off center (see the caption of Fig. 2), to avoid the nodal lines of any especially symmetric modes. We have assumed a square soundboard 1 m across (as a very rough approximation to an upright piano), with a uniform thickness of 1 cm and a density of 0.5 g/cm^3 (typical of the spruce used for soundboards). We have taken the Young's modulus to be $E=1\times10^{10}$ N/m^2 in one of the calculations, and 4×10^8 N/m^2 in the other. In both cases we used a Poisson's ratio of $\nu=0.3$. Since wood is anisotropic it has different values of these parameters along, as compared to across, the grain. These are typical along-the-grain and across-the-grain values of E for quarter cut spruce (this is the type of cut used for soundboards).[31,19] As our model assumes isotropy, the best we can do at this point is consider the behavior for these two extremal values of E. A quantity which is more difficult to estimate is the loss parameter R. Fortunately it is not as important as the others in determining the overall behavior, and especially the average magnitude, of $|Z|$. Its main role at this point is to remove singularities at the resonances and smooth out the response. Since we are interested in only the general behavior and not in the details of specific resonances, the rather coarse frequency steps used in the calculations of Fig. 2 were sufficient. The solid lines in

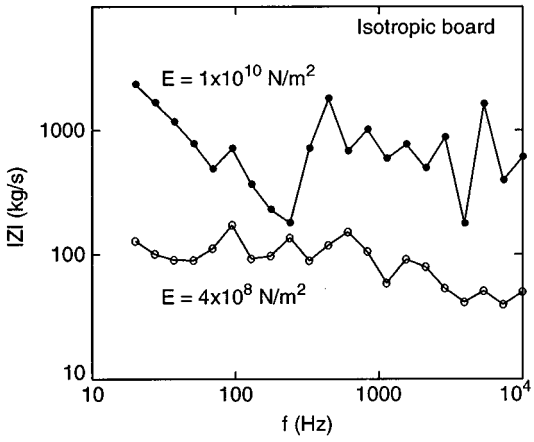


FIG. 2. Calculated driving point impedance for an elastically isotropic board. The board was 1×1 m^2, the spatial step size used in the calculation was $\Delta x=0.02$ m, and the loss factor was $R=1$ kg/(m^2−s). Filled circles: $E=1\times10^{10}$ N/m^2, $\nu=0.3$. Open circles: $E=4\times10^8$ N/m^2, $\nu=0.3$. The driving point was at $x=y=0.3$ m, with the origin at one corner of the board.

the figure simply connect the calculated values, but the overall behavior of $|Z|$ should be clear.

We first consider the behavior for the stiffer board (the larger value of E). The lowest resonant mode in this case occurs near 80 Hz; this resonance is signaled by the dip in $|Z|$ at this frequency. Additional calculations, which are not shown here, verified the location of the resonance and also its eigenvector. The resonant frequency is in agreement with the exact result for the lowest mode with these boundary conditions, which is ~ 77.5 Hz.[16] Below this frequency $|Z|$ increases as f is reduced, as expected for the stiffness dominated regime of an oscillator.[16] At higher frequencies $|Z|$ fluctuates considerably, as one passes through many resonances. The relative magnitudes of these fluctuations and the quality factors of the resonances are functions of the loss factor R, and we have chosen its value to give results which roughly match those observed for real soundboards.[13] Comparing the calculated $|Z|$ with the observations in Fig. 1 we note several differences. We begin with the result for $E=1\times10^{10}$ N/m^2, the stiffer of the two cases considered. First, the average value of $|Z|$ calculated at intermediate frequencies, $100-10^3$ Hz, is close to the experimentally measured value of 1×10^3 kg/s. However, at frequencies in the range 10^3-10^4 Hz, the calculated $|Z|$ is too high and does not fall sufficiently as frequency is increased. Indeed, it is approximately independent of frequency from below 100 to 10^4 Hz. For the softer board ($E=4\times10^8$ N/m^2) the impedance is too low at low frequencies by nearly a factor of 10, and also does not fall as rapidly at high frequencies as observed in the measurements, Fig. 1. The trends seen in going from the stiff to the soft board make it clear that no choice of E for an isotropic board will yield the behavior observed for the impedance of real soundboards.

Perhaps most troubling is the fact that the calculated impedance varies little with frequency, especially since we have argued that this is a crucial feature of the behavior of real soundboards. That this result should be found in the calculation can be understood as follows. The ''average'' value of Z will depend on the density of modes in a given frequency range. For a vibrating plate at frequencies above the lowest few modes, the density of modes is a constant,[17,16] so it is not surprising to find that $|Z|$ is also approximately constant, as we have observed. Indeed, analytic results confirm that $|Z|$ for an infinite board is a constant for this model.[32]

Since we want our model to exhibit an impedance which decreases substantially above 10^3 Hz, let us consider how one might obtain such behavior. In order to reduce $|Z|$ the velocity of the board must be larger for a given amplitude of the force; i.e., the board must be effectively softer. One way to make the board softer is to reduce the Young's modulus. However, as we have seen from Fig. 2, simply reducing E will not take us to this goal. This is not surprising in view of the density of modes argument just given. This leads us to consider a somewhat more realistic model, a board with elastic anisotropy, which will be described in the next section.

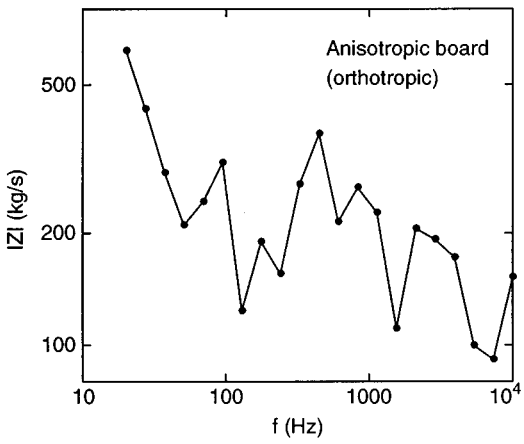


FIG. 3. Calculated driving point impedance for an orthotropic (anisotropic) board. Values of the elastic constants appropriate for spruce were used: $E_x=1\times10^{10}$ N/m^2, $E_y=4\times10^8$ N/m^2, $G=1\times10^9$ N/m^2, $\nu_x=0.4$, and $\nu_y=0.01$. The other parameters were the same as in Fig. 2.

III. ANISOTROPIC BOARD

Real soundboards are generally made of quarter cut wood,[31,19] so that elastically they are highly anisotropic. In terms of the Young's modulus the anisotropy is typically a factor of ~ 10 or more. We therefore consider an anisotropic model with different elastic constants in the x and y directions. We will imagine that the grain of the soundboard runs parallel to the x axis, so that the ''soft'' direction is along y.[33] The equation of motion for such a board, which is generally termed ''orthotropic,'' is[5,34]

$$\rho h\,\frac{\partial^2 z}{\partial t^2}=-D_x\,\frac{\partial^4 z}{\partial x^4}-(D_x\nu_y+D_y\nu_x+4D_{xy})\,\frac{\partial^4 z}{\partial x^2\,\partial y^2}-D_y\,\frac{\partial^4 z}{\partial y^4}, \tag{8}$$

where the rigidity factors are

$$D_x=\frac{h^3E_x}{12(1-\nu_x\nu_y)},\quad D_y=\frac{h^3E_y}{12(1-\nu_x\nu_y)},\quad D_{xy}=\frac{h^3G_{xy}}{12}. \tag{9}$$

Here E_x and ν_x are the Young's modulus and Poisson's ratio for the x direction, etc., for y, and G_{xy} is the shear modulus. Converting to a set of difference equations, the form is the same as (5) and the coefficients become

$$a_1=\frac{2}{1+\beta},\quad a_2=\frac{-1+\beta}{1+\beta},\quad a_3=\frac{-D_x(\Delta t)^2}{\rho h(\Delta x)^4},$$
$$a_4=\frac{-D_y(\Delta t)^2}{\rho h(\Delta x)^4},\quad a_5=\frac{-D_{xy}(\Delta t)^2}{\rho h(\Delta x)^4},\quad a_6=\frac{(\Delta t)^2}{\rho h(\Delta x)^2}, \tag{10}$$

where we have included a driving force and damping as before.

Some results for $|Z|$ for a square orthotropic plate are given in Fig. 3. Here we have used a set of elastic constants which have been measured for a real soundboard,[5] and which are also close to values given elsewhere in the literature.[31]

Other plausible values for the elastic constants yield similar results. Comparing with the experiments in Fig. 1 we find that $|Z|$ is too small by about a factor of 4 in the range $100-10^3$ Hz, implying that our orthotropic model is not stiff enough in this range. In addition, the calculated impedance again varies much too weakly with frequency above $\sim 10^3$ Hz, as it decreases by less than a factor of ~ 2 in going from 10^3 to 10^4 Hz (note that the vertical scale in Fig. 3 spans only one decade, and is thus quite different from those in our other impedance plots). The measured impedance drops by a factor of ~ 30 in the same range. This weak frequency dependence of $|Z|$ should come as no surprise, since for an orthotropic board the density of modes is again independent of frequency at high frequencies.[5]

IV. ANISOTROPIC BOARD WITH RIBS

The behavior found for the orthotropic board in Fig. 3 indicates that $|Z|$ is too small at low frequencies. We thus need to make the board stiffer in this frequency range; this can be accomplished by choosing a material with a larger Young's modulus, or employing a thicker board, or both. However, we would still have the problems of too little frequency dependence from 10^3-10^4 Hz, and an impedance which is too large at high frequencies. We need to devise a way to construct a board which is effectively stiffer at low frequencies (long wavelengths), but is at the same time softer at high frequencies (short wavelengths). From a qualitative standpoint this can be accomplished by adding ribs to the board. Consider ribs that run along the ''soft direction,'' which in our case is y; i.e., across the grain as on a real soundboard. Ribs can increase the stiffness in two ways. First, they can have a larger Young's modulus than the across-the-grain value; that is, $E_{\rm rib} > E_y$. Second, they add thickness to the board thus increasing h, and this increases the rigidity as can be seen from (9). In words, this simply means that a thick board is harder to bend than a thin one, even if E is held fixed. The ribs thus cause long wavelength vibrations to experience a stiffer board. Here ''long'' means a wavelength comparable to or greater than the rib spacing. On the other hand, vibrations with wavelengths smaller than the rib spacing can effectively fit into the spaces between ribs and thus experience a softer board characteristic of E_y.

These arguments have motivated us to study a model in which ribs are added to the orthotropic board of the previous section.[35] Ours is not the first soundboard model to include ribs, as such a model was also considered by Kindel.[4,5] However, his concern was with the precise eigenfrequencies and eigenvectors of the low-frequency modes (generally below a few hundred Hz), whose wavelengths are larger than the rib spacing. He was therefore able to employ a model similar to (8) with suitably chosen, i.e., spatially averaged, values for the rigidities D_x, D_y, and D_{xy}, and other parameters. In this way, he was able to employ an equation of motion similar to (8) which was appropriate (and accurate) for the long wavelength modes which were his primary concern. However, since we are also interested in the short wavelength modes we must explicitly include the position dependences of the elastic constants and board thickness. Our approach is motivated by the finite difference form for the equation of motion (8), which can be interpreted as simply Newton's second law for the mass elements of volume $h(\Delta x)^2$ which make up the board.

We begin with an orthotropic board similar to that treated in the previous section. We then add ribs and let them be evenly spaced and run parallel to the y direction (which is perpendicular to the direction of the ''grain'' in the model). The ribs are taken to be narrow, and for simplicity we give them a width of one spatial grid size, Δx. For our typical calculations this corresponds to 1 to 2 cm, which is close to the width in a real case.[36] In the regions of the board between the ribs [i.e., for the corresponding values of (i,j)] the equation of motion is the same as (8) with the rigidities given previously (9), since the elastic restoring forces and the geometry of the mass elements are the same as for a board without ribs. For locations on the ribs we assume that we can use an equation of motion of the same form, but with several parameters altered due to the different thickness and Young's modulus of the ribs. For the rib locations we take

$$D_{x,\rm rib}=\frac{h^3 E_x}{12(1-\nu_x\nu_y)}, \quad D_{y,\rm rib}=\frac{(h+h_{\rm rib})^3 E_{\rm rib}}{12(1-\nu_x\nu_y)}, \quad D_{xy,\rm rib}=\frac{(h+h_{\rm rib})^3 G_{xy}}{12}, \tag{11}$$

where $h_{\rm rib}$ is the thickness of a rib, and $E_{\rm rib}$ is its Young's modulus (we ignore any anisotropy here). The physical motivation for these expressions for the rigidities can be understood as follows. Since a rib is only one spatial step wide and runs parallel to y, the connection to neighboring parts of the board in the x direction will be the same as if there were no rib present. Hence the resistance to bending in the direction perpendicular to the rib, and thus $D_{x,\rm rib}$, will be the same as without ribs. Things are different along the rib direction y as here the rib must bend with the board. For this reason, the rigidity $D_{y,\rm rib}$ is determined by the thickness of the board plus rib, $h+h_{\rm rib}$, and the Young's modulus of the rib, $E_{\rm rib}$ (assumed to be much larger than E_y). For the shear term we have for simplicity assumed that it is the same as without ribs, except that we allow for the thickness of the rib.[37] Finally, we have neglected changes in Poisson's ratio on the ribs, as these have a numerically small effect in (11). Our expressions for the rigidities are at best approximate; in particular, they ignore factors which arise due to the fact that the ''neutral surface'' of a rib, i.e., the plane through the center of mass of a rib is displaced relative to that of the board without ribs.[5] Including these effects has been shown to alter the form of the equation of motion, although the additional terms appear to be numerically small in our case. In any event, to do a better job would require a more extensive finite-element analysis, with much smaller spatial step sizes (much smaller than the rib width). While such an analysis might be warranted (and would be interesting to pursue) in the future, we believe that at this stage a computationally simpler, albeit approximate, approach will yield more physical insight. A similar approach yielding a ''practical'' theory of the orthotropic plate has been shown to give results close to the exact, and much more involved, theory.[7]

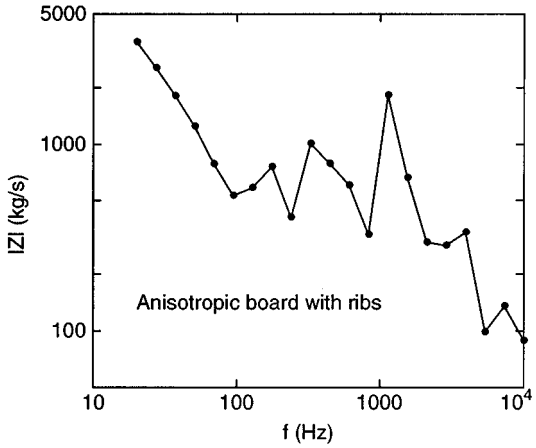


FIG. 4. Calculated driving point impedance for an orthotropic (anisotropic) board with ribs. The elastic constants were the same as those used in Fig. 3. We also took $h_{\rm rib}=0.01$ m and $E_{\rm rib}=2\times10^{10}$ N/m^2. The board had 12 evenly spaced ribs.

Our numerical scheme for a board with ribs may be summarized as follows. At all locations (i,j) we use the finite difference result (5) to compute $z(i,j,n+1)$. For locations (i,j) which are not on a rib we use the coefficients a_i from (10) with the rigidities (9). For locations which are on a rib we use the same form for the a_i coefficients, but with the following changes: The rigidities are taken from (11) and the factors of h which appear explicitly in (10) are replaced by $h+h_{\rm rib}$. The loss factor R is the same as for the off-rib locations.

We have investigated the behavior of our model with ribs for a variety of parameter sets, and the results presented in Fig. 4 are typical. The elastic constants used here are those measured by Kindel for a real soundboard, and which he found to give good agreement with experiment in his model calculations of the low-frequency modes for a grand piano. These parameter values are also close to those reported elsewhere in the literature.[31] We have assumed a rib thickness of 1 cm and a Young's modulus of the rib of $E_{\rm rib}=2\times10^{10}$ N/m^2. As one would expect from (11), very similar results are found if the rib thickness is increased and $E_{\rm rib}$ is reduced. The results in Fig. 4 exhibit all of the features that we have argued are musically important. First, the fundamental frequency is near 100 Hz. Second, $|Z|$ is approximately constant with a value near 10^3 kg/s in the range $100-10^3$ Hz. Third, the impedance drops significantly as the frequency increases from 10^3 to 10^4 Hz. All of these features agree at least semiquantitatively with measurements on real soundboards, and they can also be readily understood on a physical basis.

At low frequencies the ribs increase the stiffness of the board in the y direction. This increases the impedance at low frequencies as compared to the orthotropic board without ribs in Fig. 3. For the results shown here we assumed that there are 12 ribs, which seems to be a typical value, and that they are evenly spaced across the board. This rib configuration together with the values used for the rib thickness and Young's modulus given above, resulted in a value of $|Z|$ at 100 Hz which is quite close to the desired value. Our calculations show that this impedance can be increased by adding more ribs, increasing their thickness, increasing $E_{\rm rib}$, or any combination of these changes.

For the calculation shown in Fig. 4, the ''average'' value of the impedance (i.e., the value found after averaging over the resonances) was approximately constant from 100 up to $\sim$2000 Hz, and then decreased at higher frequencies. The frequency at which $|Z|$ began to decrease was found to depend on the number of ribs. As the number of ribs was reduced this frequency dropped, while adding ribs kept $|Z|$ constant up to a higher frequency. This agrees with the arguments given at the beginning of this section. The ribs increase the stiffness and hence increase $|Z|$ for wavelengths comparable to or longer than the rib spacing, while allowing the board to be ''softer'' at shorter length scales. Indeed, we see that the impedance at 10^4 Hz for the board with ribs approaches the value observed for the orthotropic board without ribs, Fig. 3.

The results in Fig. 4 agree fairly well, but not perfectly, with the measured results. The main difference between the two is that the calculated impedance at 10^4 Hz is somewhat larger, here by about a factor of 3–5, than the experimental value. This difference could perhaps be eliminated by fine tuning the elastic constants and adjusting the properties of the ribs (their number, spacing, length, and thickness). However, given the approximate nature of our model, we do not consider such ''fits'' to be warranted at present. It is also interesting to compare our calculations with the measurements of Conklin;[19] as mentioned above, Conklin reported that Z is essentially frequency independent from 1000–5000 Hz. It is seen that the calculated impedance for a board with ribs in Fig. 4 changes only a little over this range, with most of the decrease at higher frequencies. Hence our calculation seems consistent with the experiments of both Wogram[13] and Conklin.[19] In any event, it seems safe to conclude that this model contains the ingredients necessary to account for the general behavior observed experimentally.

V. OTHER SOUNDBOARD PROPERTIES

So far we have focused our attention entirely on the behavior of $|Z|$. In this section we will consider two other important properties, the phase associated with the impedance and the sound production. Results for the phase angle ϕ between the driving force and the soundboard velocity are shown in Fig. 5. These results are for the board with ribs considered in Fig. 4. At low frequencies, $f<100$ Hz, the board velocity is largely out of phase with the driving force, as expected for an oscillator when f is below the frequency of the fundamental mode. In the intermediate range the phase angle approaches zero, as F and v are essentially in phase with each other. Hence at these frequencies the transfer of energy from the string to the board is fairly efficient. Above about 10^3 Hz the phase angle decreases, as observed experimentally. The overall behavior is in fairly good agreement with the measurements in Fig. 1. The only significant difference appears to be at the highest frequencies, where the calculated phase angle approaches $-90°$ somewhat more slowly than found in the measurements. The reason for this is not completely clear, but it may be due to discreteness

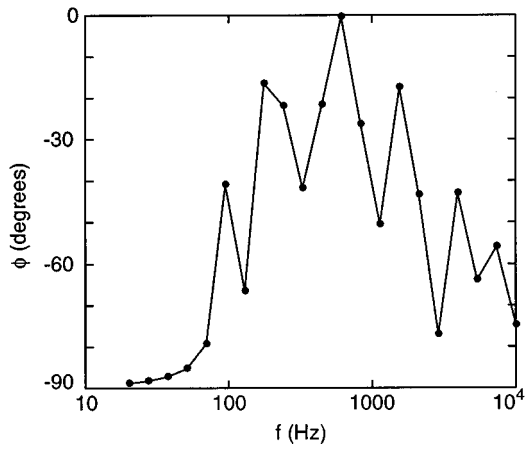


FIG. 5. Calculated phase angle between the driving force and the soundboard velocity for the soundboard with ribs considered in Fig. 4.

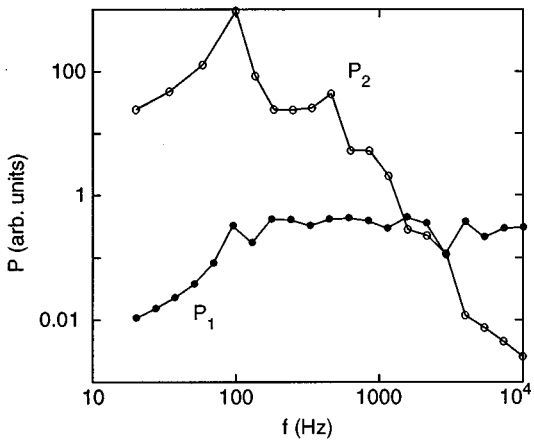


FIG. 6. Calculated sound production for the soundboard with ribs considered in Fig. 4, using the approximations (12) and (13). Note that the vertical scales are arbitrary, and are not the same in the two cases.

effects in our finite-element approach (such as taking the ribs to be one spatial unit wide), rather than being inherent in the model.

As we have mentioned several times, one motivation for our work has been the desire to construct a soundboard model which can be used for simulation studies involving string vibrations. The hope is to combine string simulations like those in Refs. 1 and 2 with our model soundboard; that is, to effectively terminate the strings in the simulations at a point on our model soundboard. This would allow one to investigate the coupling between strings and other related issues. Of course, one goal of such calculations would be to produce a convincing piano sound, and this brings up the question of how to calculate the sound produced by our theoretical soundboard.

The calculation of the sound produced by a vibrating board is not as straightforward as the calculations of the impedance and phase angle considered above. The usual treatments of sound production by vibrating sources are usually limited to geometrically simple objects, and deal only with the frequency domain, since the propagation problem (in the surrounding medium) is then simpler than when viewed in the time domain.[38] However, our square board with ribs is not a simple source and things will become more complicated when (in the future) we consider more realistic soundboard geometries. We also do not wish to work in the frequency domain, for reasons already mentioned. These considerations lead us to consider two approximate approaches. Our aim here is to judge if either might be adequate for further simulation studies.

In the first approach we assume that each finite difference element of the board contributes to the radiated power independently of all of the other elements, and that the power contributed by an element is proportional to the square of its velocity. With these assumptions the total radiated power is

$$P_1 \sim \sum_{i,j} v(i,j)^2, \tag{12}$$

where $v(i,j)$ is the velocity of element (i,j) and is calculated by averaging the instantaneous board velocity $v(i,j,n+1)=[z(i,j,n+1)-z(i,j,n)]/\Delta t$ over one period of the driving force. This approach ignores the interference of the sound generated by different regions of the board, which leads to a reduction of the radiation efficiency with increasing frequency.[16] We therefore expect that (12) will overestimate the sound level at high frequencies. Even with this significant deficiency (and we should expect other difficulties at low frequencies[38]), it is instructive to consider the behavior of P_1.

In our second approach to the calculation of the radiated sound, we treat the board as effectively a single oscillator and assume that the total radiated power is proportional to the square of the average velocity of the entire board. This approach, which has been used in guitar modeling,[10] yields the sound power

$$P_2 \sim \left[\sum_{i,j} v(i,j) \right]^2. \tag{13}$$

This expression allows in a crude sense for cancellation of sound produced by different regions. Before discussing the behavior of P_1 and P_2 we wish to emphasize that they are only *approximations*. Their ultimate usefulness can only be determined by comparison with a more detailed treatment.

Results for P_1 and P_2 for our soundboard with ribs are shown in Fig. 6. Note that we have followed the experiments[13] and plotted the radiated power for a fixed board velocity (v_0) at the driving point as a function of frequency. The expressions given above for P_1 and P_2 were evaluated by averaging over one complete period of the driving force, after the system had reached a steady state. Approximations P_1 and P_2 both exhibit a rapid increase as f increases from 20 to about 100 Hz. At higher frequencies P_1 is essentially constant up to 10^4 Hz, while P_2 exhibits a maximum at $\sim$100 Hz and a rapid decrease at higher frequencies. When comparing with the measurements in Fig. 1 we see that P_1 is much too large at high frequencies; that is, P_1 does not decrease sufficiently (or at all) at high frequencies. This deficiency can be readily understood from the arguments given above. On the other hand, we see that P_2 compares fairly well with the measured result. In particular, the calculated decrease as the frequency varies from a few

hundred Hz to 10^4 Hz is about a factor of 10^4, while the observed value is the same to within about a factor of 3. It thus appears that the approximation P_2 may be satisfactory for future simulations.

VI. DISCUSSION

To the best of our knowledge the only experimental results for the impedance and other properties of a piano soundboard over the full ''high'' frequency range (10^3–10^4 Hz) are those in Ref. 13. We have compared our model results to those measurements where possible, and have used what we have argued are the important features of the experimental results for $|Z|$ to guide us to the model with ribs. The overall agreement is encouraging as our calculated impedance has the correct overall magnitude and frequency dependence, at least semiquantitatively. Nevertheless, many refinements of the model are needed. These include the addition of the bridges, allowing motion in all three dimensions, and adding ''crown'' to the board. Further study of how best to calculate the sound produced by a vibrating soundboard is also needed.

We have argued that the ribs are essential for producing the desired variation of the impedance with frequency. It is interesting to note that Wogram has reported a few results for soundboards without ribs.[25] He measured $|Z|$ as the ribs were removed in stages by grinding them off an originally complete soundboard. Surprisingly, while removing the ribs in this way did have some effect on the frequencies of the lowest modes, the general behavior of $|Z|$, even above 10^3 Hz, was essentially the same after the ribs were completely removed. This result runs counter to our arguments. There are two ways to explain this discrepancy. (1) Our arguments are wrong. Obviously, we do not believe that this is the case. If Wogram's results for a soundboard without ribs are really correct, then the impedance of an orthotropic plate should exhibit the features seen in Fig. 1. However, given our calculations for this case we do not see how the impedance of an orthotropic plate without ribs can be made to resemble the measured result. One place for significant maneuvering in our model concerns the loss term. We have employed a very simple and numerically convenient choice for the loss, but one could imagine a rather different possibility. Suppose that the loss is dominated by sound production. We know from the experiments that the efficiency of sound generation drops substantially above 10^3 Hz. This effect could then make the board effectively softer at high frequencies and reduce $|Z|$. We do not believe that this is a plausible scenario,[39,40] but cannot discount it at present. (2) There may be a problem with the experiments. We are reluctant to comment much on this point, but do note that the complete removal of a rib may not be a simple task. If the glue used to attach the ribs to the soundboard had penetrated into the board, which seems likely, this could add extra stiffness even when the rib material itself was completely ground away. This discussion also raises the possibility that the results in Fig. 1 for a soundboard with ribs may need to be reexamined. This possibility may seem a bit discouraging from the point of view of this paper, since we have made use of trends in the measured impedance in Ref. 13 to lead us to our final soundboard model. However, the key feature we have utilized is the decrease of $|Z|$ at high frequencies, and we have given arguments independent of the experiments (involving the frequency dependence of the decay rate, as discussed above) that such behavior should be expected.[41] In any event, more experimental work would clearly be desirable.

In summary, we have investigated the properties of a simple finite-element model of a piano soundboard, and are encouraged by the level of agreement between our results and experimental measurements. We have argued that ribs are an essential ingredient in obtaining this agreement, but further experiments are needed to determine if this is in fact the case.

ACKNOWLEDGMENTS

We are grateful to P. F. Muzikar and A. J. Korty for useful comments, especially concerning the manuscript, A. Tubis for remarks concerning timpani vibrations, G. Weinreich (and also one of the referees) for some very interesting comments regarding the experimental situation, and B. Martin for inspiring discussions. We are also indebted to H. W. Jackson for providing a copy of Ref. 5. This work was supported in part by the NSF through Grant No. PHY-9722031.

[1] A. Chaigne and A. Askenfelt, ''Numerical simulations of piano strings. I. Physical model for a struck string using finite difference methods,'' J. Acoust. Soc. Am. **95**, 1112 (1994).
[2] A. Chaigne and A. Askenfelt, ''Numerical simulations of piano strings II. Comparisons with measurements and systematic exploration of some hammer-string parameters,'' J. Acoust. Soc. Am. **95**, 1631 (1994).
[3] I. Nakamura, ''The vibrational character of the piano soundboard,'' Proc. 11th ICA, Paris (1983), Vol. 4, p. 385.
[4] J. Kindel and I-C. Wang, ''Modal analysis and finite element analysis of a piano soundboard,'' in *Proceedings of the 5th International Modal Analysis Conference* (Union College, Schenectady, NY, 1987), p. 1545.
[5] J. Kindel, ''Modal analysis and finite element analysis of a piano soundboard,'' Master's thesis, University of Cincinnati, 1989.
[6] While not strictly intended to apply to piano soundboards, the work of Ref. 7 on the vibrations of orthotropic plates is also quite relevant to the problems considered in the present paper.
[7] G. W. Caldersmith, ''Vibrations of orthotropic rectangular plates,'' Acustica **56**, 144 (1984); G. Caldersmith and T. D. Rossing, ''Determination of modal coupling in vibrating rectangular plates,'' Appl. Acoust. **17**, 33 (1984).
[8] H. L. Schwab, ''Finite element analysis of a guitar soundboard,'' Catgut Acoust. Soc. **24**, 13 (1975); H. L. Schwab and K. C. Chen, ''Finite element analysis of a guitar soundboard—Part II,'' Catgut Acoust. Soc. **25**, 13 (1976).
[9] O. Christensen and B. B. Vistisen, ''Simple model for low-frequency guitar function,'' J. Acoust. Soc. Am. **68**, 758 (1980).
[10] O. Christensen, ''An oscillator model for analysis of guitar sound pressure response,'' Acustica **54**, 289 (1984).
[11] A. Chaigne, ''On the use of finite differences for musical synthesis. Application to plucked string instruments,'' J. Acoust. **5**, 181 (1992).
[12] See, for example, B. E. Richardson, G. W. Roberts, and G. P. Walker, ''Numerical modeling of two violin plates,'' J. Catgut Acoust. Soc. **47**, 12 (1987).
[13] K. Wogram, ''Acoustical research on pianos. Part I: Vibrational characteristics of the soundboard,'' Das Musikinstrument **24**, 694 (1980); **24**, 776 (1980); **24**, 872 (1980).
[14] Note that while Ref. 13 plotted the sound-pressure level, we have chosen here to plot the sound power level. This affects the vertical scale, but not the overall shape of this log-log plot.
[15] H. Suzuki, ''Vibration and sound radiation of a piano soundboard,'' J. Acoust. Soc. Am. **80**, 1573 (1986).

[16] N. H. Fletcher and T. D. Rossing, *The Physics of Musical Instruments* (Springer-Verlag, New York, 1991).
[17] A. H. Benade, *Fundamentals of Musical Acoustics* (Oxford U. P., New York, 1976).
[18] G. Weinreich, ''Coupled piano strings,'' J. Acoust. Soc. Am. **62**, 1474 (1977).
[19] H. A. Conklin, Jr., ''Design and tone in the mechanoacoustic piano. Part II. Piano structure,'' J. Acoust. Soc. Am. **100**, 695 (1996).
[20] The frequency at which this falloff begins depends somewhat on the position on the soundboard where the impedance is measured, and was observed (Ref. 13) to vary by a factor of ~ 2. However, above this point $|Z|$ decreased at approximately the same rate with frequency at all locations on the soundboard.
[21] D. W. Martin, ''Decay rates of piano tones,'' J. Acoust. Soc. Am. **19**, 535 (1947).
[22] T. C. Hundley, H. Benioff, and D. W. Martin, ''Factors contributing to the multiple rate of piano tone decay,'' J. Acoust. Soc. Am. **64**, 1303 (1978).
[23] We should also note that Ref. 19 has shown the results of impedance measurements which do not show any significant decrease in the impedance at high frequencies. However, those results extend to only about 5000 Hz, so they may be consistent with those of Wogram (Ref. 13). In any event, it would be very desirable to re-examine the experimental situation, as will be discussed further below.
[24] Conklin (Ref. 19) gives a particularly instructive discussion of soundboard design and behavior.
[25] K. Wogram, in *The Acoustics of the Piano*, edited by A. Askenfelt (Royal Swedish Academy of Music Publication No. 64, Stockholm, 1990), p. 83.
[26] L. D. Landau and E. M. Lifshitz, *Theory of Elasticity* (Pergammon, London, 1959).
[27] There is a substantial literature (see, for example, Ref. 28 and references contained therein) on the vibrations of plates whose thickness varies with position, since structures of this type serve many practical functions. In most cases the analysis of such stepped plates, which are similar in some respects to the model with ribs which we will discuss below, is limited to plates with a single step. Additional steps can be accommodated within the theory, but only with significant computational complications which end up making the problem essentially a numerical one. Given our ultimate goal of modeling a semirealistic soundboard, it therefore seems reasonable to take a numerical approach from the outset.
[28] I. E. Harik, X. Liu, and N. Balakrishnan, ''Analytic solution to free vibration of rectangular plates,'' J. Sound Vib. **153**, 51 (1992).
[29] N. Giordano, *Computational Physics* (Prentice-Hall, Upper Saddle River, NJ, 1997).
[30] Other details of the calculation, such as the choice of time step, etc., follow standard practice, as discussed in Refs. 29, 11, and 1.
[31] V. Bucur, *Acoustics of Wood* (CRC, Boca Raton, 1995).
[32] L. Cremer and M. Heckl, *Structure-Borne Sound* (Springer-Verlag, New York, 1973).
[33] For most upright pianos, including the one studied in Ref. 13 for which schematic results are given in Fig. 1, the grain of the soundboard runs diagonally, rather than parallel to one of the edges as we assume here. This is one of the details that we will assume can be neglected in a ''first'' approximation.
[34] S. G. Lekhnitskii, *Anisotropic Plates* (Gordon and Breach, New York, 1968).
[35] For simplicity we will not include the bridges in our model. Since the bridges generally run approximately parallel to the stiff direction (in our case x), we expect that they will affect the behavior much less than the ribs.
[36] Ideally one would like to have the spatial step size in a calculation be much smaller than any of the important length scales in the problem. However, using a significantly smaller value of Δx would add greatly to the computational demands, but not, we think, much to the physical insight.
[37] Our approach is in the same spirit as the approximate theory for orthotropic plates developed in Ref. 7, where it was shown that in practice, accurate results are obtained if the shear modulus is chosen such that $D_{xy} = \sqrt{D_x D_y}$. Our choice of G_{xy} here and above satisfies this criterion to within 20% or better. Variations of the value of G_{xy} by this amount did not have a significant effect on the results.
[38] P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (Princeton U. P., Princeton, NJ, 1986).
[39] It is known from Ref. 40 that air loading has a quite significant effect on the modes of a timpani membrane. However, given its different thickness and mass density, we would not expect air loading to be nearly as important for a piano soundboard.
[40] R. S. Christian, R. E. Davis, A. Tubis, C. A. Anderson, R. I. Mills, and T. D. Rossing, ''Effects of air loading on timpani membrane vibrations,'' J. Acoust. Soc. Am. **76**, 1336 (1984).
[41] It has been pointed out by G. Weinreich (private communication) that other factors connected with the decreasing length of the strings may also play a major role in producing faster decays at high frequencies. This would appear to be a (currently unsettled) question which requires more experimental work.

Detection of temporal gaps in noise in dolphins: Evoked-potential study

Vladimir V. Popov and Alexander Ya. Supin
Institute of Ecology and Evolution, Russian Academy of Sciences, 33 Leninsky Prosp., 117071 Moscow, Russia

(Received 21 June 1996; accepted for publication 2 April 1997)

Temporal resolution of hearing was studied in bottlenosed dolphins by recording the auditory brain-stem response (ABR) evoked by gap in noise. Gaps shorter than 0.5 ms evoked a response combining both off- and on-components; longer gaps evoked separate off- and on-responses. Both the response to a short gap and on-response to the end of a long gap increased with increasing gap duration. On-response recovered completely at gap duration of 5–10 ms. Small but detectable response arose at gap duration as short as 0.1 ms. Contrary to the on-response after a long silence, the response to a short gap was less dependent on noise intensity. From these data, the temporal transfer function of the supposed integrator was derived assuming nonlinear transform of the integrator output to ABR amplitude. Equivalent rectangular duration of the found temporal transfer function was 0.27 ms. © *1997 Acoustical Society of America.* [S0001-4966(97)07207-X]

PACS numbers: 43.80.Lb, 43.80.Jz, 43.64.Bt [FD]

INTRODUCTION

Many natural sounds contain rapid amplitude fluctuations. Therefore, it is of interest to know how well the auditory system responds to these fluctuations. The auditory system of echolocating dolphins is of special interest in this respect since dolphins use rapidly successive locating sound pulses of very short duration (for a review see Au, 1993).

In the past, attempts were made to estimate the temporal resolution in dolphins using behavioral methods. However, the complexity of the behavioral methods limited the data collection. Perhaps it was the reason why estimations of temporal resolution of the dolphin's hearing varied widely. Analysis of echolocation data has shown the integration time to be between 200 and 300 μs (Au *et al.*, 1988; Au, 1990). Experiments with discrimination between pulse pairs have also suggested that pulses merge into an ''acoustic whole'' when separated no longer than 200–300 μs (Dubrovskiy, 1990). Backward masking in dolphins is also possible at intervals of up to 200–300 μs (Moore *et al.*, 1984; Dubrovskiy, 1990). However, a much longer integration time (dozens of ms) has been found by Johnson (1968, 1991) using temporal summation data for tone pulses of various frequencies.

Preceding studies have shown that the recording of evoked potentials may be advantageous in investigating the unique abilities of the dolphin's auditory system. In particular, the auditory brain-stem response (ABR) may be of use for measuring a number of auditory characteristics in dolphins (Ridgway *et al.*, 1981; Popov and Supin, 1990a,b; Supin *et al.*, 1993; Popov *et al.*, 1995).

ABR recording has shown very high temporal resolution of the dolphin's hearing. In conditions of paired click stimulation, ABR recovery time in dolphins is several times shorter than in other mammals. In particular, complete recovery requires interstimulus intervals of only several ms, while a just detectable response to the second click is observed at intervals as short as 200–300 μs (Supin and Popov, 1985, 1995a; Popov and Supin, 1990a). In agreement with this result, the ABR in dolphins was shown to be able to follow rhythmic sound clicks up to rates as high as 1500–2000/s (Popov and Supin, 1990b). The ABR can also form the envelope following response (EFR) when sound amplitude fluctuates rhythmically with a rate of up to 2000 Hz (Dolphin, 1995; Supin and Popov, 1995b). As regards the temporal summation limit, ABR amplitude increases with stimulus duration of up to 300–500 μs (Popov and Supin, 1990a).

These data do indicate high temporal resolution, but they do not supply precise measurements of the temporal transfer function and integration time in the dolphin's auditory system. The motivation of the present study was to find the integration time course in dolphins. For this purpose, we used a commonly adopted method to estimate the auditory temporal resolution, namely the temporal gap technique: evaluation of the shortest detectable silent interval in an otherwise continuous sound. This method was very productive to estimate temporal resolution of the auditory system (e.g., Plomp, 1964; Penner, 1977; Smiarowski and Carhart, 1975; Fitzgibbons and Wightman, 1982; Fitzgibbons, 1983; Moore *et al.*, 1993). Our preliminary study has shown that very short gaps in noise are effective to evoke the ABR in dolphins (Supin and Popov, 1995c). In the present study, we investigated the effect in more detail using a variety of gap durations and noise intensities. Thereafter, we used the obtained data to derive the temporal transfer function and estimate the integration time in the dolphin's auditory system.

I. MATERIAL AND METHODS

A. Subjects

The experimental animals were two bottlenosed dolphins, *Tursiops truncatus*, with no apparent indication of any disease. The animals were caught 2–3 months before the study and were adapted to the keeping environment and to

the experimental procedure. They were kept at the Utrish Marine Station of the Russian Academy of Sciences (Black Sea coast) in sea-water pools. The care and use of the animals were performed under the guidelines established by the Russian Ministry of Higher Education on the use of animals in biomedical research.

B. Experimental conditions

During experimentation, a dolphin was placed in a bath ($4\times0.6\times0.6$ m) filled with seawater. The animal was supported by a stretcher so that the dorsal part of the body and blowhole were above the water surface. The animal was neither anesthetized nor curarized. The everyday experimental session lasted for 3 to 4 h, after which the animal was returned to its home pool.

C. ABR recording

The ABRs were recorded using 1-cm disk electrodes secured at the body surface with a drop of adhesive electric-conductive gel. The active electrode was placed on the dorsal part of the head, and the reference electrode near the dorsal fin, both remaining above the water surface. The recorded potentials were amplified, bandpass filtered between 200 and 5000 Hz, and averaged using the Neuropack-II evoked response recorder. To collect ABR, the recorder was triggered at the pregap burst onset, thus a whole complex of responses to the both noise bursts and the gap was recorded. Averaging of 512 to 1024 responses was enough to measure response parameters with satisfactory precision.

D. Stimuli

Stimuli were noises generated by activation of a spherical piezoceramic transducer by quasirandom binary sequence (strobe time 2 μs). The signal was gated in bursts and attenuated. Gap-containing stimuli consisted of a pre-gap burst, gap of varying duration, and post-gap burst. The pre-gap and post-gap bursts were each 5 ms long; the gap durations varied among values as follows: 0.1, 0.15, 0.2, 0.3, 0.5, 0.7, 1, 1.5, 2, 3, 5, 7, and 10 ms. In a few experiments, the 10-ms pre-gap burst was used; this difference did not influence results markedly.

The transducer has a frequency characteristic peaking at 50 kHz. This characteristic dictated the noise spectrum [Fig. 1(a)]. Because of transducer ''ringing'' and reverberation in the bath, a pause in noise could not be made shorter than 0.1 ms [Fig. 1(b)]. For data collection, stimuli were presented at a rate of 10/s.

The transducer was immersed in water at a depth of 30 cm, 0.75 m in front of the animal's head. To increase the regularity of the acoustic field and shorten reverberation, the bath walls and free water surface in front of the animal were covered with sound-absorbing material (rubber with closed air cavities). Intensity and duration of stimuli were monitored through a hydrophone located near the animal's head with a passband of 150 kHz. During noise bursts, broadband noise intensity varied from 60 to 140 dB *re*: 1 μPa of root-mean-square sound pressure as measured by the hydrophone. These intensities corresponded to 0–80 dB *re*: response threshold.

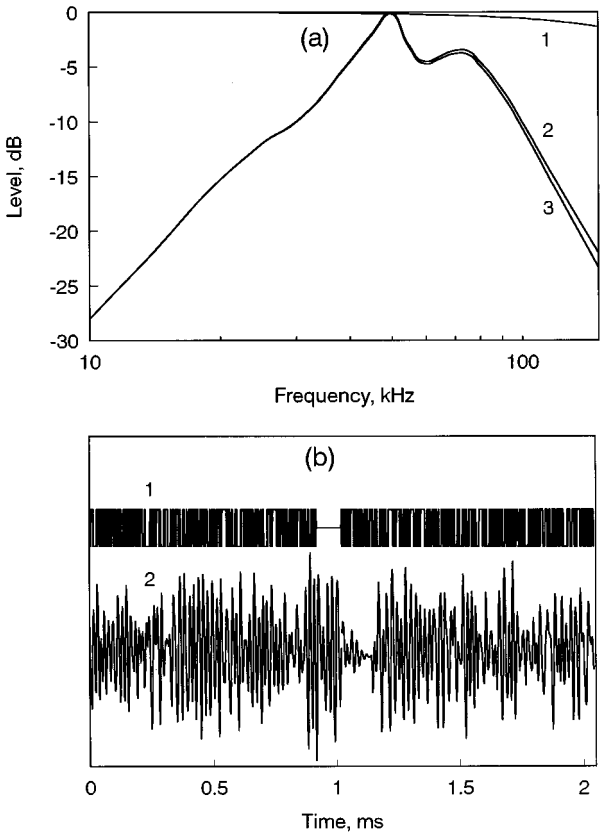


FIG. 1. Stimulus characteristics. (a) Long-term frequency spectrum of the quasi-random binary electric signal (1), the transducer frequency response (2), and resulting long-term spectrum of the acoustic signal (3). (b) A fragment of (1) electric and (2) acoustic signals containing a 0.1-ms gap.

II. RESULTS

The ABRs to noise bursts with a gap are exemplified in Fig. 2. In this experiment, the pre-gap noise burst was 5 ms long; then it was followed by a gap of varying duration, and the post-gap burst 5 ms long. All the records contain the on-response to the first burst onset and a response to the gap. The first on-response was identical in all the records. It consisted of a series of waves occurring mainly within the first 5–6 ms following the noise onset. Response amplitude was as large as a few μV when measured between the highest positive and negative peaks, as the double-headed arrow at the upper record shows.

The response to a gap depends on the gap duration. At long duration (2–10 ms), it contained separated off-response to the first burst offset (gap beginning) and on-response to the second burst onset (gap end). When gap durations were 10 ms or longer, the second on-response was equal to the first one; with shortening the gap, the response amplitude diminished. Off-responses were of lower amplitude and simpler configuration than on-responses; the latencies of the main positive and negative waves of the off-response were 0.4–0.5 ms longer than those of the on-response.

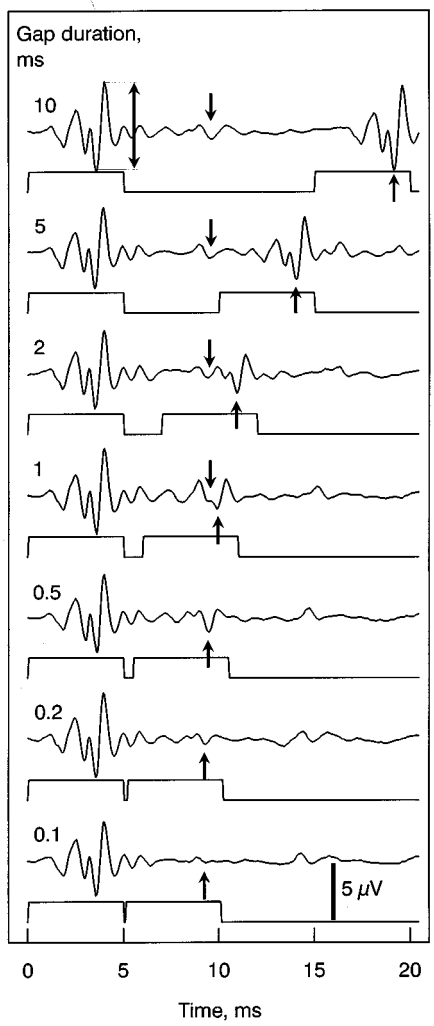


FIG. 2. The ABR evoked by two noise bursts with a gap of various duration between them. In each pair of records, the upper one is the evoked-potential record (negativity upward) and the lower one is the stimulus record (noise onset upward). Sound intensity 100 dB *re*: 1 μPa. Gap durations (ms) are indicated near the curves. Double-headed arrow at the upper record shows the response amplitude as it was measured in this study. At long gaps (1–10 ms), downward directed arrows show off-responses (to the gap beginning) and upward directed arrows show on-responses (to the gap end); at short gaps (0.1–0.5 ms), gap-responses (fused off- and on-responses) are shown by arrows.

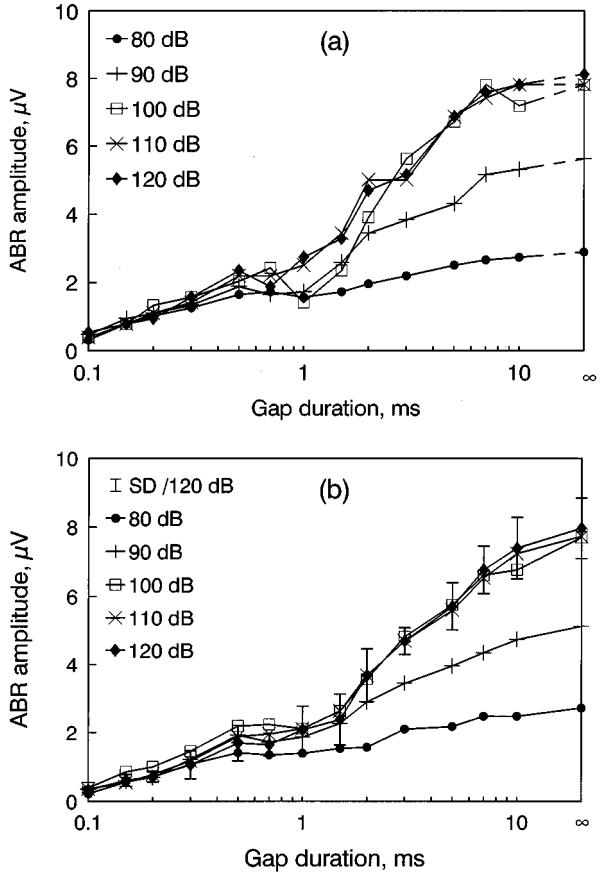


FIG. 3. ABR amplitude dependence on gap duration at sound intensities of 80–120 dB *re*: 1 μPa root-mean-square, as indicated. (a) Representative example of data from one recording series. (b) Average data of five recording series in two animals (means; s.d. are shown for the series of 120 dB only). Point ''∞'' at the abscissa indicates response amplitude after a long silence.

With shortening the gap, the delay between the off-response to the gap beginning and on-response to the gap end became respectively shorter. At gap durations shorter than 1 ms, the two responses fused into a single response which is reasonable to designate as the gap-response. Since the off-response had a longer latency than the on-response, at the 0.5-ms gap duration they overlapped almost completely. This combined gap-response to the 0.5-ms gap was of a higher amplitude than the on-response (the more so than the off-response) at longer gap durations (1 ms in Fig. 2). Further shortening of the gap resulted in diminishing of the response. However, small but detectable response was observed at the shortest available gap duration, i.e., at 0.1 ms.

Thus, below, we use the term ''on-response'' for gap durations of 1 ms and longer when this response can be distinguished from the off-response, and the term ''gap-response'' for gap durations 0.5 ms and shorter.

The obtained data are presented in Fig. 3(a) and (b) as plots showing the dependence of the gap- and on-response amplitude on gap duration. Figure 3(a) exemplifies data of one representative recording session; Fig. 3(b) presents averaged data (means ±s.d.) of five recording sessions in two animals. The functions were obtained at several noise intensities: from 80 to 120 dB *re*: 1 μPa (20–60 dB above the response threshold). Standard deviations in Fig. 3(b) are shown only for one noise intensity (120 dB) to make the figure hardly readable. At gap durations from 0.1 to 0.5 ms, the combined gap-response amplitude was plotted; at longer durations, the on-response amplitude was plotted. The plots show the main effects as were exemplified by original records in Fig. 2: Response amplitude increased with gap prolongation, except that the gap-response to the 0.5-ms gap was higher than the on-response to 0.7 to 1 ms gaps.

Another feature of the plots in Fig. 3 is that they group densely at gap durations from 0.1 to 0.5 ms, although they become widely separated at longer durations. In other words, response amplitude at short gap durations was almost independent of the noise intensity, although response to the noise onset after a long silence did depend, in large measure, on intensity. This effect is also demonstrated by original records in Fig. 4 which show on-responses to noise onset and re-

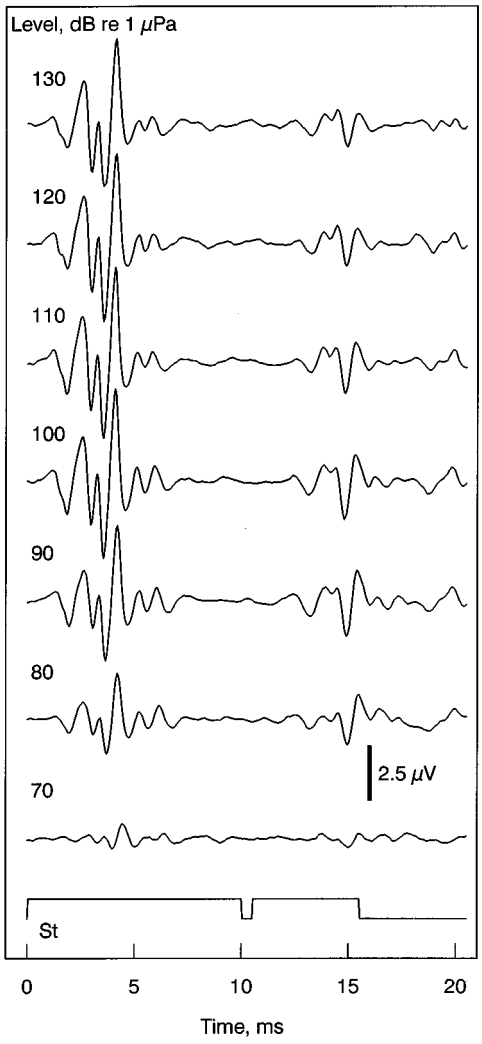


FIG. 4. ABR evoked by noise with a 0.5-ms gap at various sound intensities (as indicated near the records). St—stimulus record.

sponses to 0.5-ms gap at various noise intensities. The records show that both responses rose proportionally to the intensity only at very low intensities (70–80 dB *re*: 1 μPa, i.e., 10–20 dB above the response threshold). Further increase of the noise intensity (90–120 dB) resulted in increasing the first on-response amplitude; however, the gap-response amplitude remained almost constant (90–100 dB) or even diminished (110–120 dB).

This effect is demonstrated quantitatively by plots showing response amplitudes versus noise intensity at various gap durations (Fig. 5). At gap durations from 0.2 to 0.5 ms, the combined gap-response amplitude was plotted; at longer durations, the on-response amplitude was plotted. The plots show that after a long silence (∞), on-response amplitude was intensity dependent until 100 dB *re*: 1 μPa (40 dB above the response threshold), beyond which it reached asymptote. In this region, the dependence was roughly close to a straight line when the intensity was expressed on a dB (logarithmic) scale and the amplitude on a linear scale. Response to a short gap was intensity dependent only in a limited range beyond which it remained at a constant level or even diminished at higher intensities. The shorter gap duration, the shorter was the intensity range where response amplitude was intensity dependent.

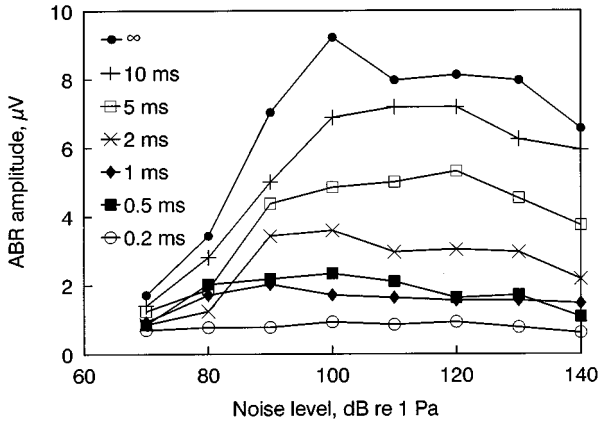


FIG. 5. Gap-evoked ABR amplitude dependence on sound intensity at various gap durations (as indicated). Sign ''∞'' indicates response after a long silence.

III. DISCUSSION

A. Computational of the temporal transfer function

Temporal resolution of the auditory system may be expressed in terms of its temporal transfer function which has integration properties: the shorter the integration time, the better the temporal resolution. Thus measurement of temporal resolution means, first of all, determination of the integrating temporal transfer function.

Consider the events at the integrator input and output when presenting a sound burst with a gap [Fig. 6(a)]. If a constant input signal I appears at the integrator input at an instant t_0 (burst onset), then the output signal is

$$R(t)=I\int \varphi(t-t_0)dt, \tag{1}$$

where t is time, $R(t)$is the integrator output (response), and $\varphi(t)$ is the sought-for temporal transfer function. If the burst is long enough, the integrator output reaches the asymptote:

$$R_{\max}=I\int_0^{\infty}\varphi(t)dt. \tag{2}$$

If the signal disappears at the integrator input at an instant t_1 (a gap beginning), the output signal is

$$R(t)=R_{\max}-I\int \varphi(t-t_1)dt. \tag{3}$$

The second sound onset at an instant t_2 (the end of the gap) restores the integrator output to the previous level of $R_{\max}$. Thus the response to the gap is evoked by a shift of the integrator output from $R_{\max}$ to $R(t)$ (off-response), or from $R(t)$ to $R_{\max}$ (on-response), or both (gap-response).

However, the ABR amplitude is not proportional to this shift since it depends nonlinearly on signal intensity. To demonstrate the effect of the nonlinear transform, Fig. 6(b) shows the same functions as in Fig. 6(a) but presented in a logarithmic ordinate scale. Such presentation, in particular, explains why the gap-response is weakly dependent on

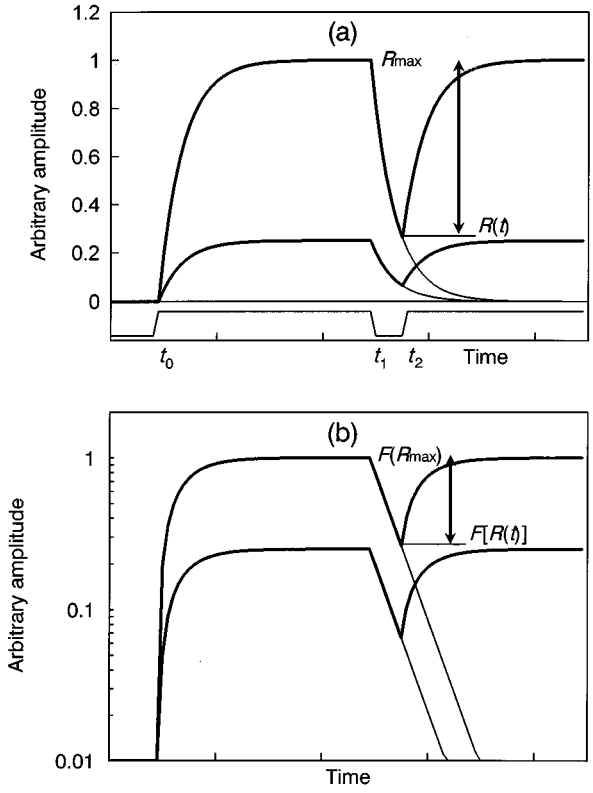


FIG. 6. Diagrammatic illustration of nonlinear transform of the integrator output during the gap-in-noise stimulation. (a) Linear presentation of the integrator output. Two plots show integrator outputs as two different sound intensities. R_{max} and $R(t)$—integrator output levels at the gap beginning and end, respectively (shown for higher intensity integrator output only); the double-headed arrow shows the gap-response amplitude. t_0, t_1, and t_2—the first burst onset, first burst offset (gap beginning), and second burst onset (gap end), respectively. Thin lines after t_2 show integrator output at further gap prolongation. (b) The same functions as in (a) presented in the logarithmic ordinate scale. $F(R_{max})$ and $F[R(t)]$—logarithmic transforms of values R_{max} and $R(t)$, respectively.

sound intensity. Curves 1 and 2 show integrator output signals at two input signal levels; presented in the linear scale, curve 2 in Fig. 6(a) reaches a lower level and demonstrates a lesser gap-response than curve 1. However, after the logarithmic transform [Fig. 6(b)], curves 1 and 2 differ only in their shift along the ordinate scale, while their gap-responses become equal. When the gap is long enough for the integrator output to reach the response threshold level, the gap-response becomes intensity dependent: the longer the gap duration, the wider the intensity range where response is intensity dependent (see Fig. 5).

The intensity-to-amplitude transfer function makes it possible to find the integrator output signal $R(t)$ using ABR data. As Fig. 6(b) shows, the integrator output shift from R_{max} to $R(t)$ or back results in the response amplitude

$$A(t)=F(R_{max})-F[R(t)], \tag{4}$$

where $A(t)$ is the response amplitude when gap duration is t, and $F(R)$ is the intensity-to-amplitude transfer function. In Fig. 6, the logarithmic function is denoted as $F(R)$. Actually, the function $F(R)$is expected to reproduce the ABR amplitude dependence on stimulus intensity; i.e., the function which is presented by the plot marked ''∞'' in Fig. 5, with sound intensity expressed in the power domain (this function is not precisely logarithmic). Of course, we do not know the actual values of the integrator output $R(t)$. However, according to the adopted model, the auditory system contains a square-law rectifier followed by a temporal integrator (Veimeister, 1979; Moore *et al.*, 1988; Plack and Moore, 1990). Thus the integrator input and output signals are expected to be proportional to the sound power. So we use sound power (expressed in Pa^2) as a relative measure of I and R. It follows from Eq. (4):

$$R(t)=f[A_{max}-A(t)], \tag{5}$$

where $R(t)$ (Pa^2) is the integrator output at the time t, $A_{max}=F(R_{max})$ is the response amplitude to the given sound intensity after a long silence, and $f(A)$ ($Pa^2/\mu V$) is the function inverse to the intensity-to-amplitude transfer function $F(R)$ ($\mu V/Pa^2$). Then the sought-for temporal transfer function $\varphi(t)$ can be found since it follows from Eq. (3) that

$$\varphi(t)=-dR(t+t_1)/I\ dt. \tag{6}$$

Since both $R(t)$ and I are expressed in Pa^2, $\varphi(t)$ has dimensionality of 1/s.

Successive stages of calculations according to Eqs. (5) and (6) are illustrated in Fig. 7. Panel (a) shows $A_{max}-A(t)$ values, found using experimental data of Fig. 3(b) as the $A(t)$ function. Panel (b) shows $f[A_{max}-A(t)]$ values calculated using the plot ''∞'' of Fig. 4, as the $F(R)$function. In so doing, we had to make one additional assumption to solve the problem as follows. As was pointed out, the gap-response combining both on- and off-components is of higher amplitude than the on-response only; i.e., intensity-to-amplitude transfer functions $F(R)$ are different for these two response types. It results in breaks of plots between 0.5 and 1 ms in Fig. 7(a). Therefore, at short gaps we cannot calculate correctly $f[A_{max}-A(t)]$ values, since A_{max} is the on-response amplitude and $A(t)$ is the gap-response amplitude. In order to solve the problem, we supposed that $F(R)$ function for gap-responses is approximately 1.5 times steeper than that for on-responses. This eliminates the breaks in plots of Fig. 7(b). Figure 7(c) shows $\varphi(t)$ as a result of differentiation and dividing by I of functions presented in Fig. 7(b). Note that the functions were calculated from the time of 0.05 ms, although experimental data were obtained only for gap durations of 0.1 ms and longer. We calculated these values between $t=0$ and $t=0.1$ ms supposing $A(t=0)=0$.

The functions $\varphi(t)$ calculated at several sound intensities are very similar. The functions demonstrate a decline to the 0.5 level at a time of about 0.18 ms. Then a further long decay appears with a slope of -3.5 to -4 in the double logarithmic scale; i.e., the function is inversely proportional to the 3.5th to 4th power of time. Initial part of these functions are better demonstrated when presented in the linear scales [Fig. 8(a)].

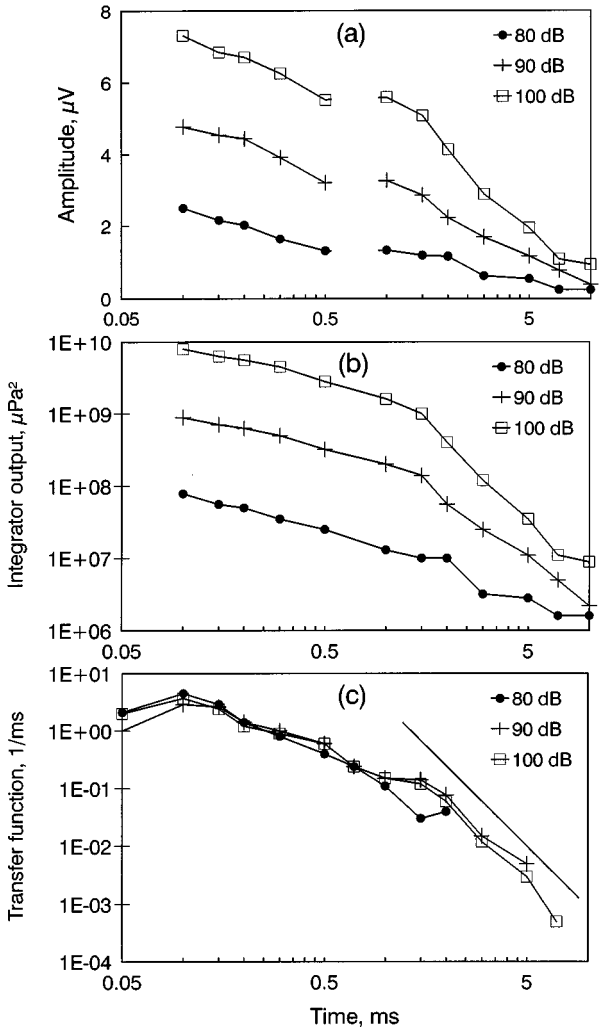


FIG. 7. Successive stage of calculation of the integrator transfer function. (a) A_{max}—$A(t)$ values [see Eq. (5)] versus gap duration for sound intensities 80–110 dB re: 1 μPa, as indicated; the calculation done using data of Fig. 3. At gap durations of 1 ms and longer, on-response amplitude is used as $A(t)$; at gap durations of 0.5 ms and shorter, gap-response amplitude is used; on-response amplitude after a long silence is used as A_{max}. (b) $f[A_{max}-A(t)]$ values versus gap duration obtained by transform of the plots (a) using the plot ''∞'' of Fig. 5 as the $f(A)$ function. (c) Integrator transfer function obtained by differentiation and dividing by intensity of plots (b); oblique straight line shows the slope of $t^{-3.5}$.

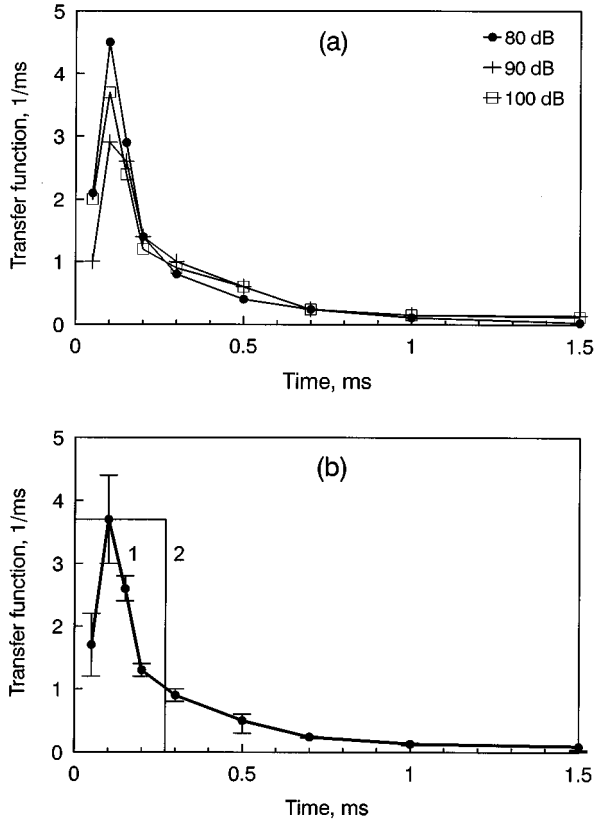


FIG. 8. (a) Temporal transfer function presented in linear scales, obtained using data at sound intensities from 80 to 110 dB re: 1 μPa, as indicated. (b) Averaged temporal transfer function, mean ±s.d. (1) and its equivalent rectangular transfer function (2).

The equivalent rectangular duration (ERD) is a commonly adopted metric of temporal resolution (Plack and Moore, 1990). Figure 8(b) shows the averaged temporal transfer function obtained as means of the plots in Fig. 8(a) and its equivalent rectangular function. ERD was found to be 0.27 ms.

Similar temporal transfer function was found in dolphins using the double-click test (Supin and Popov, 1995a). There is some difference between the ERD of the temporal transfer function found by the double-click test (Supin and Popov, 1995a) and by the gap test (this study): 0.35 and 0.27 ms, respectively. However, the differentiation procedure used in the present study is very error-sensitive because it involves the subtractions of values that do not differ much. Taking this into account, the agreement between the previous and present data may be considered satisfactory.

B. Comparison of temporal resolution in dolphins with those in other animals and humans

Exact comparison of the present data with those in other animals and humans is difficult since most of gap-detection data were obtained psychophysically while our data were obtained by the evoked-potential method. Nevertheless, preliminary comparison seems to be reasonable, taking into account that behavioral data on the integration time in dolphins (Au *et al.*, 1988; Au, 1990; Dubrovskiy, 1990) agree well with the present and previous (Supin and Popov, 1995a) evoked-potential data.

It appears that the dolphin's hearing is sensitive to extremely short gap durations as compared to other animals and humans. Detectable ABR is evoked by a gap as short as 0.1 ms. So it is reasonable to assume that the gap detection threshold in dolphins is not more than 0.1 ms. In humans, gap-in-noise detection thresholds revealed by psychophysical measurements were 3–5 ms, i.e., more than an order of magnitude longer than in dolphins (Plomp, 1964; Penner, 1977; Smiarowski and Carhart, 1975; Fitzgibbons and Wightman, 1982; Fitzgibbons, 1983). Recent measurements have shown that at the best combinations of the noise bandwidth and upper cutoff frequency, gap thresholds in humans may be as short as 2.2 ms (Snell *et al.*, 1994); however, this value is also an order of magnitude longer than in dolphins. Gap thresholds in sinusoids are 6–8 ms at frequencies of 400 Hz and above (Moore *et al.*, 1993). Experiments with detection

of a short stimulus in a gap of a masker (''temporal window'' measurements) have also shown the ''temporal window'' of several ms (Moore *et al.*, 1988; Plack and Moore, 1990). In several animal species, gap thresholds are of the same order, i.e., several ms (for a review see, Fay, 1988, 1992).

The difference between the dolphin and other animals is also obvious when using ERD for comparison. In humans, no method showed ERD shorter than a few ms. Careful ''temporal window'' measurements have also shown ERD of several ms (Moore *et al.*, 1988; Plack and Moore, 1990), which is more than an order of magnitude longer than ERD in dolphins.

Supposedly, the extremely high temporal resolution in dolphin is associated with the wide frequency range (more than 100 kHz) of their hearing. Indeed, many studies pointed out that gap detection thresholds decrease as the signal frequency increases (Viemeister, 1979; Fitzgibbons, 1983; Shailer and Moore, 1983; Formby and Muir, 1988). It seemed expectable since at higher frequencies, peripheral auditory filters have wider passbands, thus transferring more rapid temporal modulations. These data have been reevaluated because in those experiments the noise bandwidth was covaried with center frequency, and this was a confounding factor since fluctuations inherent in noise depend on the bandwidth, thus influencing the gap detection efficiency. Indeed, it was shown that gap sensitivity is markedly affected by the inherent fluctuations (Glasberg and Moore, 1992; Snell, 1995). Keeping the bandwidth constant, temporal resolution was less dependent on sound frequency (Shailer and Moore, 1985, 1987; Moore and Glasberg, 1988; Grose *et al.*, 1989; Plack and Moore, 1990; Eddins *et al.*, 1992; Moore *et al.*, 1993). This behavior is consistent with the argument that the width of the auditory filters have little influence on overall integration time which is dictated mostly by a central integrator. Later it was shown that both the sound frequency and bandwidth influence the gap detection (Snell *et al.*, 1994). Therefore, upper cutoff frequency determines gap sensitivity to a large extent because of widening of peripheral filters or decreasing intrinsic noise fluctuations. The dolphin's auditory system can operate with frequencies much higher than those for human's hearing. At these high frequencies, peripheral auditory filters can combine high relative frequency tuning with wide passband. Frequency tuning in dolphins is much better than found in many other mammals and humans: The Q_{10} index is around 20 (Supin *et al.*, 1993; Popov *et al.*, 1995). However, at a frequency of 100 kHz, filters of this tuning have passbands of $100/20=5$ kHz (at the -10-dB level), thus capable of transferring amplitude modulation of the same rate, i.e., with a duration of fractions of a ms. For comparison; at the frequency of 1 kHz, a filter of even less tuning ($Q_{10}=10$) has a bandwidth of only 100 Hz and is capable of transferring amplitude modulations lasting at least several ms. Even if a central integrator rather than peripheral filters plays a primary role in temporal resolution, its integration time may be evolutionary adjusted to the temporal resolution of peripheral filters. Thus the integration time has to be very short in dolphins.

ACKNOWLEDGMENTS

This study was supported by the Russian Foundation for Basic Research, Grant Nos. 94-04-11844 and 97-04-49024. Valuable comments of reviewers are greatly appreciated.

Au, W. W. L. (**1990**). ''Target detection in noise by echolocating dolphins,'' in *Sensory Abilities of Cetaceans. Laboratory and Field Evidence*, edited by J. A. Thomas and R. A. Kastelein (Plenum, New York), pp. 203–216.

Au, W. W. L. (**1993**). *The Sonar of Dolphins* (Springer-Verlag, New York).

Au, W. W. L., Moore, P. W. B., and Pawloski, D. A. (**1988**). ''Detection of complex echoes in noise by an echolocating dolphin,'' J. Acoust. Soc. Am. **83,** 662–668.

Dolphin, W. F. (**1995**). ''Steady-state auditory-evoked potentials in three cetacean species elicited using amplitude-modulated stimuli,'' in *Sensory Systems of Aquatic Mammals*, edited by R. A. Kastelein, J. A. Thomas, and P. E. Nachtigall (De Spill, Woerden, The Netherlands), pp. 25–47.

Dubrovskiy, N. A. (**1990**). ''On the two auditory systems in dolphins,'' in *Sensory Abilities of Cetaceans. Laboratory and Field Evidence*, edited by J. A. Thomas and R. A. Kastelein (Plenum, New York), pp. 233–254.

Eddins, D. A., Hall, H. W., and Grose, J. H. (**1992**). ''Detection of temporal gaps as a function of frequency region and absolute noise bandwidth,'' J. Acoust. Soc. Am. **91,** 1069–1077.

Fay, R. F. (**1988**). *Hearing in Vertebrates: A Psychophysics Databook* (Hill-Fay, Vinnetka, IL).

Fay, R. R. (**1992**). ''Structure and function in sound discrimination among vertebrates,'' in *The Evolutionary Biology of Hearing* (Springer-Verlag, New York), pp. 229–263.

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., and Wightman, F. L. (**1982**). ''Gap detection in normal and hearing-impaired listeners,'' J. Acoust. Soc. Am. **72,** 761–765.

Formby, C., and Muir, K. (**1988**). ''Modulation and gap detection for broadband and filtered noise signals,'' J. Acoust. Soc. Am. **84,** 545–550.

Glasberg, B. R., and Moore, B. C. J. (**1992**). ''Effect of envelope fluctuations on gap detection,'' Hearing Res. **64,** 81–92.

Grose, J. G., Eddins, D. A., and Hall, J. W. (**1989**). ''Gap detection as a function of stimulus bandwidth with fixed high-frequency cutoff in normal-hearing and hearing-impaired listeners,'' J. Acoust. Soc. Am. **86,** 1747–1755.

Johnson, C. S. (**1968**). ''Relation between absolute threshold and duration of tone pulses in the bottlenose porpoise,'' J. Acoust. Soc. Am. **43,** 757–763.

Johnson, C. S. (**1991**). ''Hearing thresholds for periodic 60-kHz tone pulses in the beluga whale,'' J. Acoust. Soc. Am. **89,** 2996–3001.

Moore, B. C. J., and Glasberg, B. R. (**1988**). ''Gap detection with sinusoids and noise in normal, impaired and electrically stimulated ears,'' J. Acoust. Soc. Am. **83,** 1093–1101.

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., Peters, R. W., and Glasberg, B. R. (**1993**). ''Detection of temporal gaps in sinusoids: Effects of frequency and level,'' J. Acoust. Soc. Am. **93,** 1563–1570.

Moore, P. W. B., Hall, R. W., Friedl, W. A., and Nachtigall, P. E. (**1984**). ''The critical interval in dolphin echolocation: What is it?,'' J. Acoust. Soc. Am. **87,** 2178–2187.

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.

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.

Plomp, R. (**1964**). ''Rate of decay of auditory sensation,'' J. Acoust. Soc. Am. **36,** 277–282.

Popov, V. V., and Supin, A. Ya. (**1990a**). ''Auditory brain stem responses in characterization of dolphin hearing,'' J. Comp. Physiol. A **166,** 385–393.

Popov, V. V., and Supin, A. Ya. (**1990b**). ''Electrophysiological studies of hearing in some cetaceans and a manatee,'' in *Sensory Abilities of Cetaceans. Laboratory and Field Evidence*, edited by J. A. Thomas and R. A. Kastelein (Plenum, New York), pp. 405–415.

Popov, V. V., Supin, A. Ya., and Klishin, V. O. (**1995**). Frequency tuning curves of the dolphin's hearing: Envelope-following response study,'' J. Comp. Physiol. A **178,** 571–578.

Ridgway, S. H., Bullock, T. H., Carder, D. A., Seeley, R. L., Woods, D., and

Galambos, H. (**1981**). ''Auditory brainstem response in dolphin,'' Proc. Natl. Acad. Sci. USA **78,** 1943–1947.
Shailer, M. J., and Moore, B. C. J. (**1985**). ''Detection of temporal gaps in band-limited noise: Effects of variations in bandwidth and signal-to-noise ratio,'' J. Acoust. Soc. Am. **77,** 635–639.
Shailer, M. J., and Moore, B. C. J. (**1987**). ''Gap detection and the auditory filter: Phase effect using sinusoidal stimuli,'' J. Acoust. Soc. Am. **81,** 1110–1117.
Smiarowski, R. A., and Carhart, R. (**1975**). ''Relations among temporary resolution, forward masking and simultaneous masking,'' J. Acoust. Soc. Am. **57,** 1169–1174.
Snell, K. B. (**1995**). ''The effect of sinusoidal amplitude modulation on gap detection in noise,'' J. Acoust. Soc. Am. **98,** 1799–1802.
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.
Supin, A. Ya., and Popov, V. V. (**1985**). ''Recovery cycles of the dolphin's brain stem responses to paired acoustic stimuli,'' (in Russian) Dokl. Akad. Nauk SSSR **283,** 740–743.
Supin, A. Ya., and Popov, V. V. (**1995a**). ''Temporal resolution in the dolphin's auditory system revealed by double click evoked potential study,'' J. Acoust. Soc. Am. **97,** 2586–2593.
Supin, A. Ya., and Popov, V. V. (**1995b**). ''Envelope-following response and modulation transfer function in the dolphin's auditory system,'' Hearing Res. **92,** 38–46.
Supin, A. Ya., and Popov, V. V. (**1995c**). ''Frequency tuning and temporal resolution in dolphins,'' in *Sensory Systems of Aquatic Mammals*, edited by R. A. Kastelein, J. A. Thomas, and P. E. Nachtigall (De spill, Woerden, The Netherlands), pp. 95–110.
Supin, A. Ya., Popov, V. V., and Klishin, V. O. (**1993**). ''ABR Frequency tuning curves in dolphins,'' J. Comp. Physiol. A **173,** 649–656.
Viemeister, N. F. (**1979**). ''Temporal modulation transfer function based on modulation thresholds,'' J. Acoust. Soc. Am. **66,** 1346–1380.

Simulation of ultrasonic pulse propagation through the abdominal wall

T. Douglas Mast,[a] Laura M. Hinkelman,[b] and Michael J. Orr
Department of Electrical Engineering, University of Rochester, Rochester, New York 14627

Victor W. Sparrow
Graduate Program in Acoustics, The Pennsylvania State University, University Park, Pennsylvania 16802

Robert C. Waag
Departments of Electrical Engineering and Radiology, University of Rochester, Rochester, New York 14627

(Received 1 August 1996; accepted for publication 19 March 1997)

Ultrasonic pulse propagation through the human abdominal wall has been simulated using a model for two-dimensional propagation through anatomically realistic tissue cross sections. The time-domain equations for wave propagation in a medium of variable sound speed and density were discretized to obtain a set of coupled finite-difference equations. These difference equations were solved numerically using a two-step MacCormack scheme that is fourth-order accurate in space and second-order accurate in time. The inhomogeneous tissue of the abdominal wall was represented by two-dimensional matrices of sound speed and density values. These values were determined by processing scanned images of abdominal wall cross sections stained to identify connective tissue, muscle, and fat, each of which was assumed to have a constant sound speed and density. The computational configuration was chosen to simulate that of wavefront distortion measurements performed on the same specimens. Qualitative agreement was found between those measurements and the results of the present computations, indicating that the computational model correctly depicts the salient characteristics of ultrasonic wavefront distortion *in vivo*. However, quantitative agreement was limited by the two-dimensionality of the computation and the absence of detailed tissue microstructure. Calculations performed using an asymptotic straight-ray approximation showed good agreement with time-shift aberrations predicted by the full-wave method, but did not explain the amplitude fluctuations and waveform distortion found in the experiments and the full-wave calculations. Visualization of computed wave propagation within tissue cross sections suggests that amplitude fluctuations and waveform distortion observed in ultrasonic propagation through the abdominal wall are associated with scattering from internal inhomogeneities such as septa within the subcutaneous fat. These observations, as well as statistical analysis of computed and observed amplitude fluctuations, suggest that weak fluctuation models do not fully describe ultrasonic wavefront distortion caused by the abdominal wall. [S0001-4966(97)00308-1]

PACS numbers: 43.80.Cs, 43.20.Fn, 43.58.Ta [FD]

INTRODUCTION

Much has recently been written on the distortion of ultrasonic wavefronts by tissue inhomogeneities and its effect on ultrasonic images. Direct measurements of the ultrasonic distortion produced by human abdominal wall,[1,2] chest wall,[3] and breast[4–6] have been made and techniques for the correction of this distortion have been proposed and examined.[7–15] However, the physical causes of ultrasonic wavefront distortion by human soft tissues are not yet well understood.

Several investigators have set out to improve this understanding by devising models of human tissue to explain observed distortions. Robinson *et al.*[16] and Sauerbrei[17] were able to explain shadowing, enhancement, and double image artifacts seen in abdominal imaging via ray tracing through arrangements of homogeneous structures, each with a different characteristic sound speed and a simple geometric shape. More recently, Manry and Broschat[18] applied a finite-difference time-domain (FDTD) algorithm to a similarly simple model to study ultrasonic propagation through the breast. A Dutch group[19] has developed a method to calculate acoustic transmission and reflection at an irregularly shaped boundary between layers of two homogeneous media. However, none of these models takes into account the detailed structure of human tissues or the complex arrangement of these tissues in the human body. One early study used power spectra of sectional images to determine scattered power of porcine liver tissue under the Born approximation.[20] A recent study of the effect of tissue microstructure on ultrasonic imaging has been performed by a group at the Riverside Research Institute[21] using sound-speed maps determined from acoustic microscopy images of liver cross sections. Simulations of a-scan and b-scan mode imaging were performed by convolving idealized pulses with estimated tissue impulse-

[a]Current affiliation: Applied Research Laboratory, The Pennsylvania State University, University Park, PA 16802.
[b]Current affiliation: Department of Meteorology, The Pennsylvania State University, University Park, PA 16802.

response functions. While this approach provides insight into the relationship between tissue microstructure and speckle, the linear convolution process does not model distortion effects such as beam and focus degradation.

Models of ultrasonic propagation through distorting layers, whether stated explicitly or implicitly, are essential to all distortion correction methods. Many use models of the types described above. For example, Kossoff *et al.*[7] have been able to reduce gross artifacts caused by refractive effects in the abdomen. Smith *et al.*[8] found that distortions created by imaging through the skull could be corrected by compensating for refraction by a plane layer, but had little success when viewing the abdominal wall as a plane layer of fat. Others followed the lead of those working on distortion encountered in astronomy and modelled the distorting tissue as a phase screen at the receiving aperture,[9–11] which implies that received waveforms differ only in phase. One notable exception is an early paper by Hirama *et al.*,[12] who used a phase and amplitude screen at or away from the aperture. Substantial recent experimental evidence has confirmed that the distortion produced by actual tissues is more complicated than that produced by a single phase screen at the receiving aperture.[2,3,5,6,22] For this reason, some investigators[13,14] have employed a phase screen placed some distance from the aperture, so that amplitude and wave shape variations as well as arrival time differences can be accounted for. Others[15] have approximated this configuration using both a phase and an amplitude screen at the aperture.

Models of wavefront distortion are also important because they are used to test distortion correction algorithms. Very few of these techniques have been tested using ultrasonic signals recorded after propagation through actual human tissues.[7,8,13] Instead, most investigators have relied on computer simulations in which phase and/or amplitude distortion is numerically added to received or calculated waveforms.[9–12,15] Others have used data from experiments in which an aberrator constructed of a uniform medium with varying thickness is inserted between the ultrasonic transducer and the target.[9,10,12,14]

While some of the proposed algorithms perform well under these simplified conditions, none has been able to return an ultrasonic beam or image distorted by human tissues to diffraction-limited quality. Such focus correction is theoretically possible; for instance, a wavefront emitted by a point source and distorted by propagation through an inhomogeneous medium is optimally refocused by propagating the time-reversed wavefront back through the same inhomogeneity.[14] The limitations of current methods may, in part, be due to the fact that each method rests on unrealistic assumptions about the nature of distortions produced clinically. For example, human tissues are not completely homogeneous, organs rarely occur in simple geometric shapes, and the thickness of the abdominal wall, chest wall, or intervening tissue of the breast is generally a significant fraction of the transducer focal length. A better understanding of the composition and structure of the body wall, breast, or other distorting tissues and their interaction with ultrasound would clearly aid the development of aberration correction techniques.

The purpose of the present study is to simulate ultrasonic propagation through the abdominal wall using a realistic model of tissue structure and a computational model that incorporates all wave effects such as single and multiple scattering, reflection, and refraction. The model is shown to produce distortion similar to experimental measurements. Results were also obtained using an asymptotic straight-ray approximation. Examination of the detailed wave propagation computed with the finite-difference model provides previously unavailable insight into the physical nature of ultrasonic wavefront distortion. The results suggest that simple phase screen models can explain some of the time-shift aberrations caused by the human abdominal wall, but that consideration of strong scattering effects is necessary to explain experimentally measured amplitude and waveform distortion.

I. THEORY

Ultrasonic pulse propagation through the human abdominal wall was modelled using the equations of motion for a lossless fluid with variable sound speed and density. The tissue was assumed to be motionless except for small acoustic perturbations. For such a fluid, the linearized equations of mass conservation, momentum conservation, and state are respectively

$$\frac{\partial \rho'(\mathbf{r},t)}{\partial t} + \boldsymbol{\nabla}\cdot(\rho(\mathbf{r})\mathbf{v}(\mathbf{r},t)) = 0, \tag{1}$$

$$\rho(\mathbf{r})\frac{\partial \mathbf{v}(\mathbf{r},t)}{\partial t} = -\boldsymbol{\nabla} p(\mathbf{r},t), \tag{2}$$

$$\frac{\partial p(\mathbf{r},t)}{\partial t} = c(\mathbf{r})^2\left(\frac{\partial \rho'(\mathbf{r},t)}{\partial t} + \mathbf{v}(\mathbf{r},t)\cdot\boldsymbol{\nabla}\rho(\mathbf{r})\right), \tag{3}$$

where $\rho(\mathbf{r})$ and $c(\mathbf{r})$ are the spatially dependent ambient density and sound speed, $\rho'(\mathbf{r},t)$ is the acoustic perturbation in density, $p(\mathbf{r},t)$ is the acoustic pressure, and $\mathbf{v}(\mathbf{r},t)$ is the (vector) acoustic particle velocity. The linear propagation equations (1)–(3) are obtained from the full fluid-mechanical equations by removing all terms of quadratic or higher order in the acoustic perturbation variables ρ', p, and $\mathbf{v}$.[23]

Equation (3) may be used to eliminate the acoustic density perturbation from Eqs. (1) and (2). This yields, in two-dimensional Cartesian coordinates, the coupled equations

$$\frac{\partial p(x,y,t)}{\partial t} + \rho(x,y)c(x,y)^2\boldsymbol{\nabla}\cdot\mathbf{v}(x,y,t) = 0, \tag{4}$$

$$\rho(x,y)\frac{\partial \mathbf{v}(x,y,t)}{\partial t} + \boldsymbol{\nabla} p(x,y,t) = 0. \tag{5}$$

Equations (4) and (5) were solved numerically using a finite-difference time-domain (FDTD) method. In order to implement the finite-difference algorithm, the fluid-mechanical equations were written in the form

$$\frac{\partial \mathbf{S}(x,y,t)}{\partial t} + \frac{\partial \mathbf{F}(\mathbf{S}(x,y,t))}{\partial x} + \frac{\partial \mathbf{G}(\mathbf{S}(x,y,t))}{\partial y} = 0, \tag{6}$$

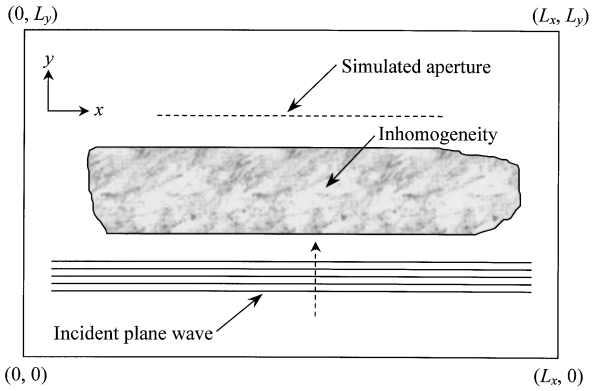


FIG. 1. Computational domain for finite-difference time-domain calculations. A plane wave pulse propagates in the y direction through water, shown as the white background, and an inhomogeneous region, shown here as a textured object. The acoustic pressure is recorded by a number of simulated transducer elements beyond the inhomogeneity.

where the ordered triplets $\mathbf{S}(x,y,t)$, $\mathbf{F}$, and $\mathbf{G}$ are defined

$$\mathbf{S}=\left(\frac{p}{\rho c^2},\rho u,\rho v\right),\quad \mathbf{F}=(u,p,0),\quad \mathbf{G}=(v,0,p), \tag{7}$$

and u and v, respectively, are the **x** and **y** components of the acoustic velocity.

The system of equations expressed in Eq. (6) was numerically solved using a two-step MacCormack algorithm that was fourth-order accurate in space and second-order accurate in time.[24,25] The implementation of the finite-difference operators was equivalent to that described in Ref. 26. The initial condition was chosen to simply model the experimental condition of a slowly-varying, nearly planar wavefront emitted from a wide band, pulsed, point-like source far from the tissue layer. The initial wavefront was represented as a plane wave pulse propagating in the $+y$ direction,

$$\begin{aligned} &p(x,y,0)=-\sin(k_0(y-y_0))e^{-(y-y_0)^2/(2\sigma^2)},\\ &u(x,y,0)=0,\quad v(x,y,0)=\frac{p(x,y,0)}{\rho c}, \end{aligned} \tag{8}$$

where the wave number k_0 is equal to $2\pi f_0/c$ for a center frequency of f_0 and σ is the Gaussian parameter of the pulse temporal envelope. The Gaussian parameter σ was chosen to simulate the bandwidth of the pulse used in the experiments, as discussed below in Sec. II.

The computational configuration is sketched in Fig. 1. 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. Periodic boundary conditions were applied on the edges of the grid parallel to the direction of propagation, that is,

$$\mathbf{S}(0,y,t)=\mathbf{S}(L_x,y,t). \tag{9}$$

The periodic boundary conditions, together with the plane wave initial condition, ensured that the wavefront remained undistorted in the absence of propagation-path inhomogeneities. On the edges normal to the direction of propagation, radiation boundary conditions were applied to calculate p and v. These conditions were chosen to absorb waves normally incident to the boundaries $y=0$ and $y=L_y$, so that waves propagating in directions close to normal would incur only small reflections. The radiation conditions, equivalent to unidirectional wave equations applied at the top and bottom boundaries, were

$$\begin{aligned} &\frac{\partial p}{\partial t}=c\frac{\partial p}{\partial y},\quad \frac{\partial v}{\partial t}=c\frac{\partial v}{\partial y},\quad \text{for } y=L_{yi},\\ &\frac{\partial p}{\partial t}=-c\frac{\partial p}{\partial y},\quad \frac{\partial v}{\partial t}=-c\frac{\partial v}{\partial y},\quad \text{for } y=0. \end{aligned} \tag{10}$$

The derivatives in Eqs. (10) were calculated using first-order-accurate differences, and the estimated time derivatives of p and v were used to advance the solution on the top and bottom boundaries after differencing was performed on the interior of the grid.

II. METHOD

The accuracy of the finite-difference method in this study was tested using a benchmark scattering computation. The benchmark problem, which was chosen to approximate the realistic problem of scattering from a single fat lobule while allowing the possibility for comparison with a known exact solution, was the scattering of a single-frequency, 3.75-MHz plane wave by a cylinder of diameter 4.0 mm with sound speed and density values equal to 95% of background. These parameters correspond approximately to the contrast and size of a typical fat lobule in a water background. The finite-difference computation was implemented using the methods detailed above, with a plane-wave radiation condition [Eqs. (10)] on the top boundary and periodic boundary conditions on the side boundaries. The incident wave was generated by the oscillating-wall boundary condition

$$v(x,0,t)=\sin(2\pi f_0 t), \tag{11}$$

where f_0 was 3.75 MHz.

The total pressure field obtained using the finite-difference method was compared with the total pressure calculated from an exact solution for scattering from the same cylinder.[27] The pressure was compared using a simulated aperture of 154 point receivers extending 12 mm in the x direction and located 1 mm above the cylinder boundary in the y direction. As seen in Fig. 2, the amplitude and phase of the pressure calculated with the finite-difference method agree very well with those predicted by the exact solution. This agreement is quantified by the L^2 error[28] between the total pressure from the finite-difference solution and the exact solution, defined as

$$\epsilon=\sqrt{\left(\sum |\hat{p}_{\text{fd}}-\hat{p}_{\text{exact}}|^2\right)\Big/\left(\sum |\hat{p}_{\text{exact}}|^2\right)}, \tag{12}$$

where $\hat{p}_{\text{fd}}$ and $\hat{p}_{\text{exact}}$ are Fourier components of the pressure calculated from the finite-difference algorithm and the exact solution, respectively. The agreement achieved corresponds to an L^2 error $\epsilon=0.015$.

Five fresh unfixed abdominal wall sections, each from a different donor, were obtained from autopsy and stored at $-20\,^{\circ}$C. Each specimen was later thawed by immersion in room-temperature saline solution, and a grid with 1-cm spac-

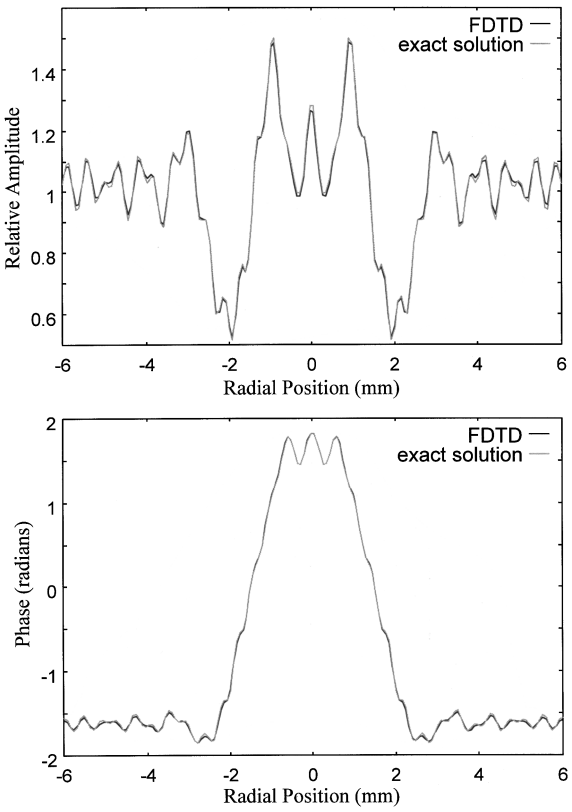


FIG. 2. Computed amplitude and phase for transmission through a fat-mimicking cylinder. Results are shown for an exact solution and for the finite-difference, time-domain (FDTD) solution. Top: theoretical and computed amplitude. Bottom: theoretical and computed phase.

ing oriented parallel to the array and elevation directions was ruled on its skin surface with India ink. The wavefront distortion produced by each specimen was then measured according to the procedure detailed in Ref. 2. A specimen was suspended between 7.5-μm-thick polyimide membranes in a water tank electronically maintained at 37.0 °C. Ultrasonic pulses at a nominal center frequency of 3.75 MHz were emitted from a custom-made, 13-mm-diam, hemispheric source[29] and, after travelling a distance of approximately 165 mm, entered the specimen through its peritoneal membrane. Data were recorded on each of the 128 elements of a 3.75-MHz linear array[30] placed 5–10 mm above the specimen's skin surface. A foam mask was used to reduce the elevation dimension of the receiving array, so that the active area of each element measured 0.72×1.44 mm^2. The array was physically translated 32 times in the elevation direction to obtain data for a 92.16×46.08 mm^2 aperture. The position of each measurement relative to the grid on the specimen's skin was noted for future correlation with the tissue cross sections used in the simulations, and the specimen was then refrozen at -20 °C.

After the wavefront distortion measurements, a bandsaw was used to slice the frozen specimens lengthwise, i.e., in the measurement array direction, along the grid rulings to obtain cross sections at 1-cm intervals. Distortion or tearing of the tissue specimens was prevented by lowering the temperature of the specimens to about -80 °C by immersion in an ethanol–dry ice bath before cutting. The specimen slices were then fixed in a 10% buffered formalin solution. The cross sections were stained with a modified Gomori's trichrome stain according to the procedure detailed in Ref. 31 so that the fat, muscle, and connective tissue could be distinguished. This stain colored the muscle red and the connective tissue blue while leaving the fat its natural color. 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 digital flatbed scanner. A commercial image editing package[32] was used to remove scanning artifacts such as shadows surrounding the specimens and to correct staining irregularities before the fat, muscle, connective tissue, and background were converted to uniform shades of green, red, blue, and white, respectively, by a FORTRAN program. The finished images were cropped to a length of about 110 mm to leave some margin beyond the size of the receiving aperture while minimizing the computational grid size. An additional guard band of about 10 mm of water was also added to each end of the images to prevent spurious waves (caused by wrap-around associated with the periodic boundary conditions on the sides of the domain) from affecting the signals recorded at the simulated aperture.

The density and sound speed arrays needed for the finite-difference computation were created by mapping the colors of the tissue images to reference density and sound speed values for the three tissue types and water. The water sound speed and density employed were those of pure water at body temperature (37.0 °C).[34,35] Sound speeds for muscle and fat were obtained by averaging values for human tissues given in Refs. 37 and 38. A representative sound speed for connective tissue was determined using an empirical formula relating collagen content to ultrasonic sound speed[39] together with a measured value for collagen content of human skin.[40] Density values for the tissues were determined from Ref. 41 by averaging values reported for adipose tissue, skeletal muscle, and skin, respectively. Attenuation values employed in the straight-ray computations were determined from measurements summarized in Ref. 37 for human fat at 37 °C, human biscep muscle at 37 °C, and human skin at 40 °C. Attenuation values reported at other ultrasonic frequencies were interpolated to the center frequency of 3.75 MHz assuming a linear dependence of attenuation on frequency. The attenuation for water was estimated by extrapolating frequency- and temperature-dependent attenuation summarized in Ref. 36 to 3.75 MHz and 37.0 °C. The parameter values employed in the present study are summarized in Table I.

The finite-difference program was used to simulate propagation of a plane wave pulse through each scanned cross section from the peritoneal membrane to the skin surface as in the distortion measurements. The spatial step size of the finite difference grid was chosen to be 0.0271 mm, or 1/15 wavelength in water at the center frequency of 3.75 MHz. The temporal step size was chosen for an optimal Courant–Friedrichs–Levy number $c\Delta t/\Delta x$ of 0.25.[33] The Gaussian parameter σ of the source pulse was chosen to be 0.3574 in accordance with the experimentally measured pulse bandwidth of 1.6 MHz. Visual comparison confirmed

TABLE I. Reference values for sound speed, density, and attenuation. Sound speed and density values were used in the finite-difference time-domain computation, while sound speed and attenuation values were used in the straight-ray computation.

Medium	Sound speed (mm/μs)	Density (g/cm^3)	Attenuation (dB/cm)
Water	1.524[a]	0.993[b]	0.02[c]
Fat	1.478[d,e]	0.950[g]	1.8[d]
Muscle	1.547[d]	1.050[g]	4.1[d]
Skin/CT	1.613[f]	1.120[g]	5.9[d]

[a]Reference 34.
[b]Reference 35.
[c]Reference 36.
[d]Reference 37.
[e]Reference 38.
[f]References 39 and 40.
[g]Reference 41.

that the simulated pulse closely matched the experimental pulses in shape and length.

Each simulation was performed on an individual IBM SP2 node with 1 GB of random-access memory. Finite-difference grids on the order of 5000×2000 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 0.444 μs, for later visualization. Computational time for each calculation was on the order of 15 hours.

Signals were recorded for 7.3 μs at a sampling frequency of 225 MHz by 128 simulated receivers of width 0.72 mm, placed about 8 mm from the skin surface of the specimen. The simulation of receiving elements was performed by integrating the locally computed pressure over a finite width. Element directivity was included by the implicit integration of omnidirectional sensitivity functions over the width of each element.

A one-dimensional version of the reference waveform method[13] was used to calculate the arrival time of the pulse at each receiving position in the simulation data. The arrival time fluctuations across the receiving aperture caused by each cross section were calculated by subtracting a linear fit from these calculated arrival times. Energy level fluctuations in the data were calculated by summing the squared amplitudes of each waveform over a 2.4 μs window which isolated the main pulse, converting to decibel units, and subtracting the best linear fit from the resulting values. 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 defined in Ref. 13.

The measured data were first corrected for gross variations in arrival time caused by the measurement geometry by subtracting a fitted two-dimensional fourth-order polynomial from the measured arrival times. Waveforms measured for the elevations which corresponded most closely to the positions from which the cross sections were taken were then analyzed using the one-dimensional technique described above. The simulation and measurement results were compared to determine the accuracy of the computational model.

Arrival time and energy level fluctuations were also computed for the modelled cross sections using simpler techniques. In each case, ultrasonic rays were assumed to pass straight through the specimen without deviation from their direction of incidence. The arrival time T of each ray was calculated using the formula

$$T(x)=\int \frac{1}{c(x,y)}dy, \tag{13}$$

where the integral is performed numerically along the ray path through the tissue and water. Likewise, the energy level of each ray was computed by integrating the spatially dependent attenuation coefficient along the ray paths, so that the transmission loss in dB was given by

$$\mathrm{TL}(x)=\int \alpha(x,y)dy, \tag{14}$$

where α is the local attenuation coefficient at the center frequency of 3.75 MHz specified in dB per unit length. The sound-speed and attenuation values employed were those reported in Table I. The resulting arrival time and energy level variations were compared to the values obtained for the same cross sections using the FDTD simulation and from the experimental measurements described above.

III. RESULTS

Tissue maps for the six abdominal wall cross sections studied are presented in Fig. 3. The average thickness of the cross sections is 26.8 mm. In each case, connective tissue layers are visible at each tissue interface, the lobular structure of the fat layer is clearly shown, and the detailed intermingling of fat and connective tissue in the muscle layer is evident. The muscles that occur in these cross sections are the rectus abdominus, transversus abdominus, and external and internal obliques. Aponeuroses, or the ends of muscle sections, are evident in each image.

Example wave fields calculated using the finite-difference model are shown in Figs. 4 and 5. These two figures show acoustic pressure fields within the tissue cross section depicted in Fig. 3(f). Examination of the wavefront evolution shows the mechanism for formation of specific features that appear in the received wavefronts shown in Fig. 6.

Figure 4 shows simulated internal wavefronts within cross section 120fe during the initial propagation through muscle and connective tissue. The distinctive feature of this portion of the propagation is a large-scale time-shift fluctuation caused by propagation through a fatty aponeurosis. Since the fat contained within the aponeurosis has a lower sound speed than the surrounding muscle and connective tissue, propagation through this region causes a substantial delay in this portion of the wavefront. This delay is accumulated as the wave propagates through the aponeurosis, so that a portion of the wavefront is delayed by about 0.2 μs (about three fourths of the wave period) as shown in Fig. 4(d). The accumulated delay is also visible in the wavefront after propagation through the entire specimen, as shown in Figs. 6(f) and 7(f). Thus, the time-shift aberration occurring in this

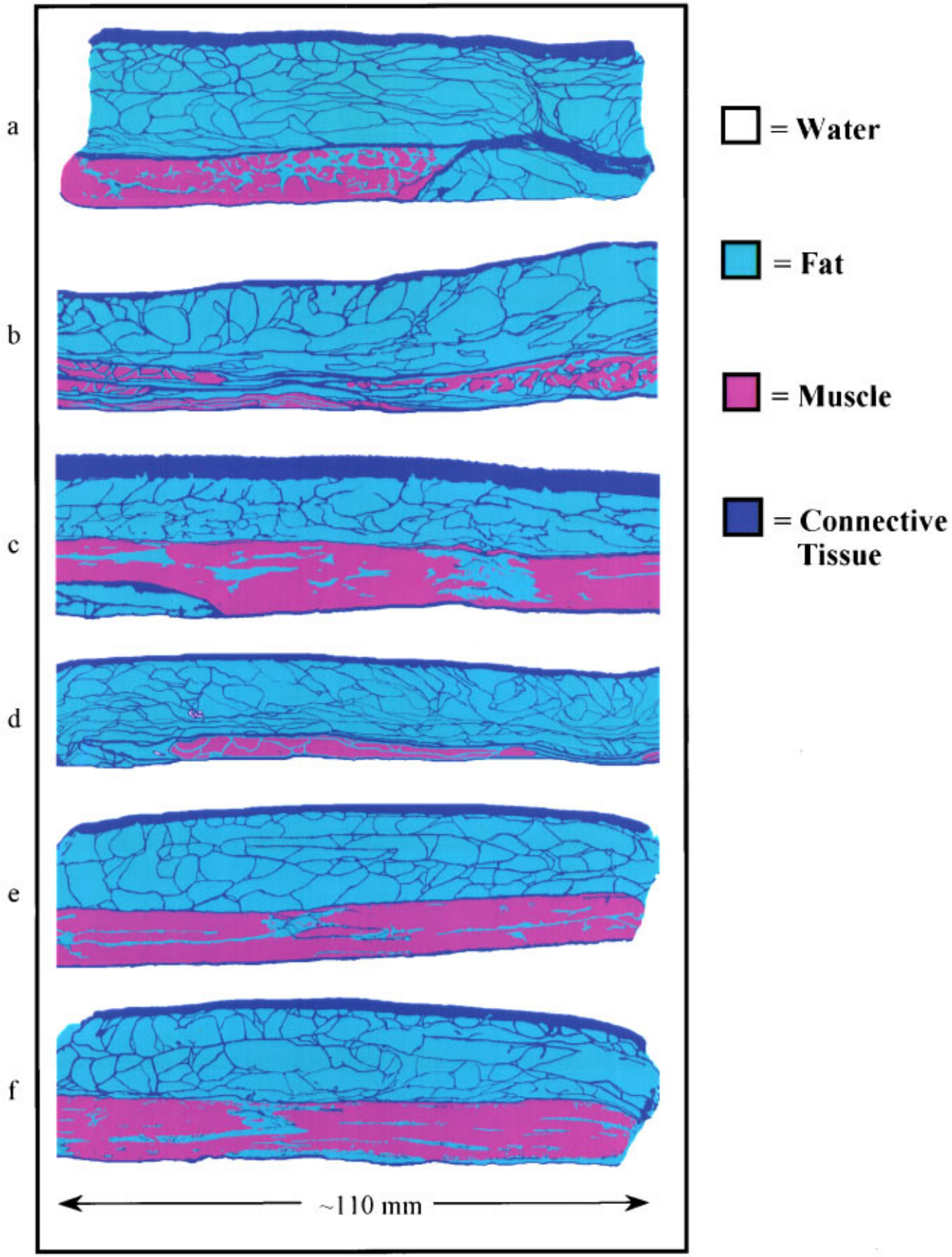


FIG. 3. Cross-sectional tissue maps. (a) 75hi. (b) 77ba. (c) 87de. (d) 102gh. (e) 120de. (f) 120fe. The number in each identifier refers to the donor of the specimen while the letters indicate the position of the cross section in the specimen.

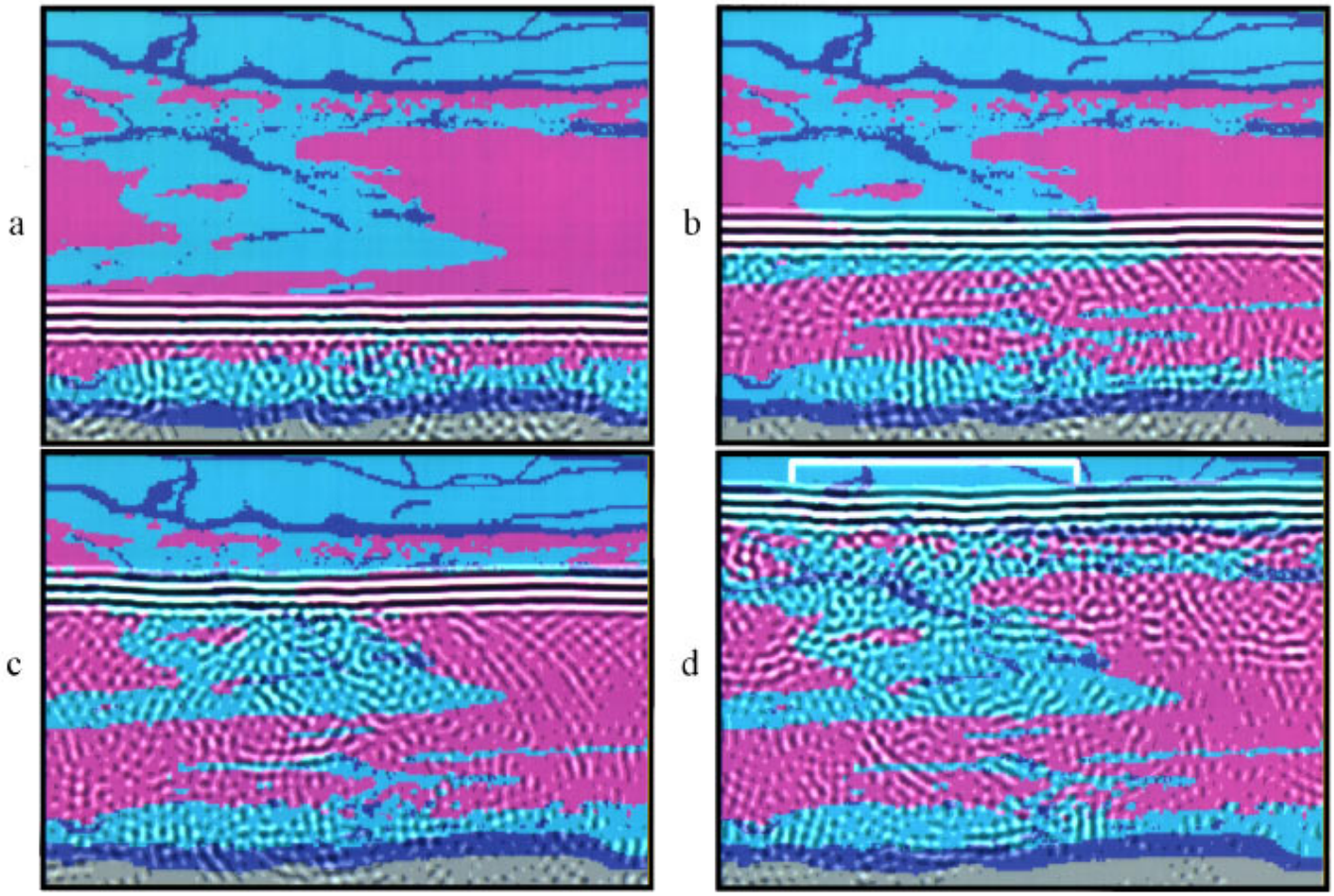


FIG. 4. Propagation through an aponeurosis in cross section 120fe. Panels (a)–(d) show the progression of the main wavefront through the muscle layer including an aponeurosis comprised of fat and connective tissue, resulting in time-shift aberration across the wavefront. The area shown in each frame is 16.0 mm in height and 18.7 mm in width. The temporal interval between frames is 1.7 μs. Tissue is shown using the same color scheme as in Fig. 3 while gray background represents water. Wavefronts are shown on a bipolar logarithmic scale with a 30 dB dynamic range. White represents maximum positive pressure and black represents maximum negative pressure. A cumulative delay of about 0.2 μs, associated with propagation through the aponeurosis, is indicated by the square bracket in panel (d).

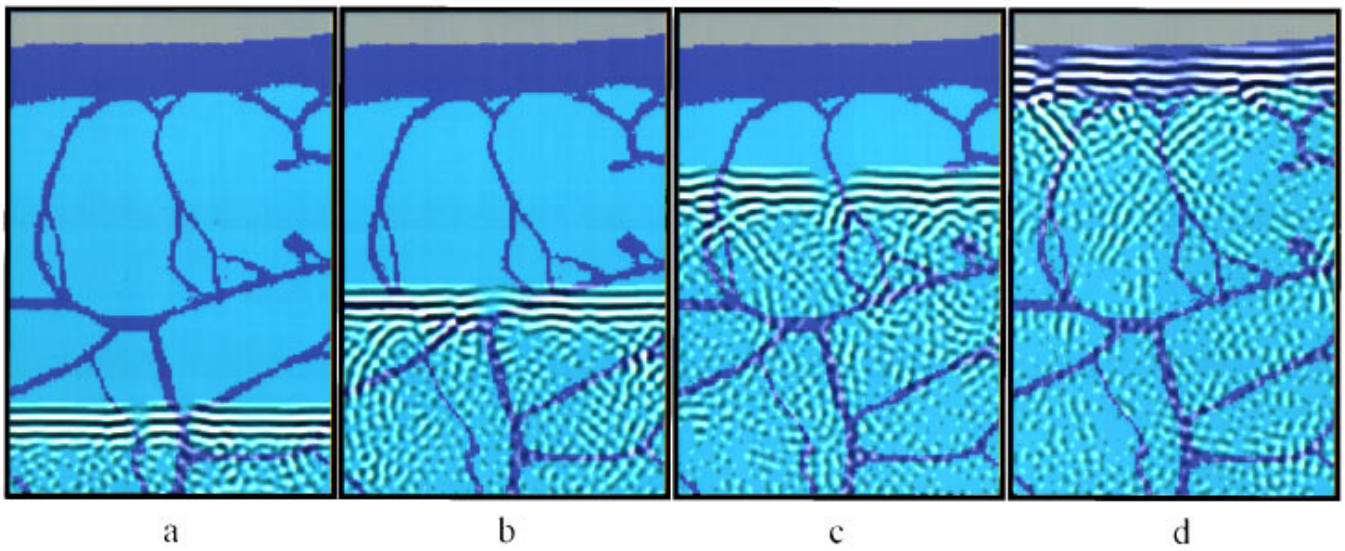


FIG. 5. Propagation through fat and septa in cross section 120fe. Panels (a)–(d) show the progression of the main wavefront through subcutaneous fat, showing the formation of amplitude dropouts by scattering from thin, near-vertically aligned septa. The area shown in each frame is 22.9 mm in height and 14.4 mm in width. The temporal interval between frames is 3.8 μs. The background medium and wavefronts are represented in the same manner as in Fig. 4.

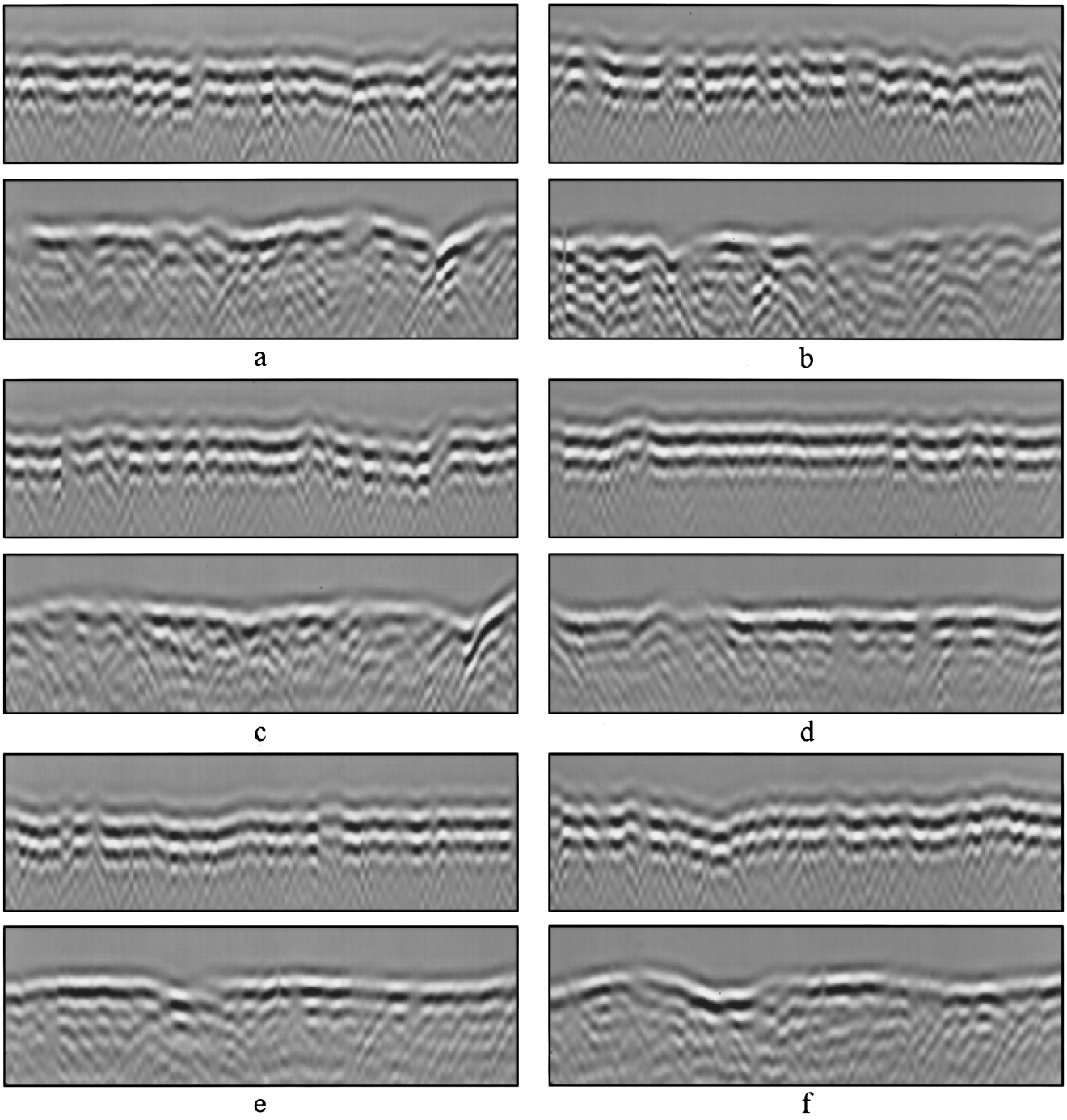


FIG. 6. Received wavefronts. In each pair, the top wavefront is the result of the FDTD propagation simulation for the given tissue map while the bottom is the corresponding experimentally measured wavefront. Wavefronts are shown on a linear gray scale with time as the vertical axis and element number as the horizontal axis. The temporal range shown is 2.0 μs for 128 elements. (a) Specimen 75hi. (b) Specimen 77ba. (c) Specimen 87de. (d) Specimen 102gh. (e) Specimen 120de. (f) Specimen 120fe.

example is associated with large-scale variations in sound speed rather than with other effects such as irregularity of the interfaces between tissue layers.

An example of amplitude dropout formation is shown in Fig. 5 using simulated internal wavefronts within cross section 120fe during propagation through the subcutaneous fat and septa. These results show the cumulative formation of amplitude fluctuations by scattering from septa aligned close to the main direction of propagation. As the wave propagates along a path that includes septa, energy is scattered outside the main direction of propagation. Scattered energy interferes constructively in a manner determined by the angle and contrast of the septa, so that secondary wavefronts are formed. Because the transmitted pulse is of short temporal duration (wide band), scattering from septa results in the removal of energy from the main wavefront, causing amplitude dropouts in the received waveforms. This effect is inherently different from effects of coherent interference that may occur for narrow band propagation through tissue.

Animated visualization of the saved pressure raster im-

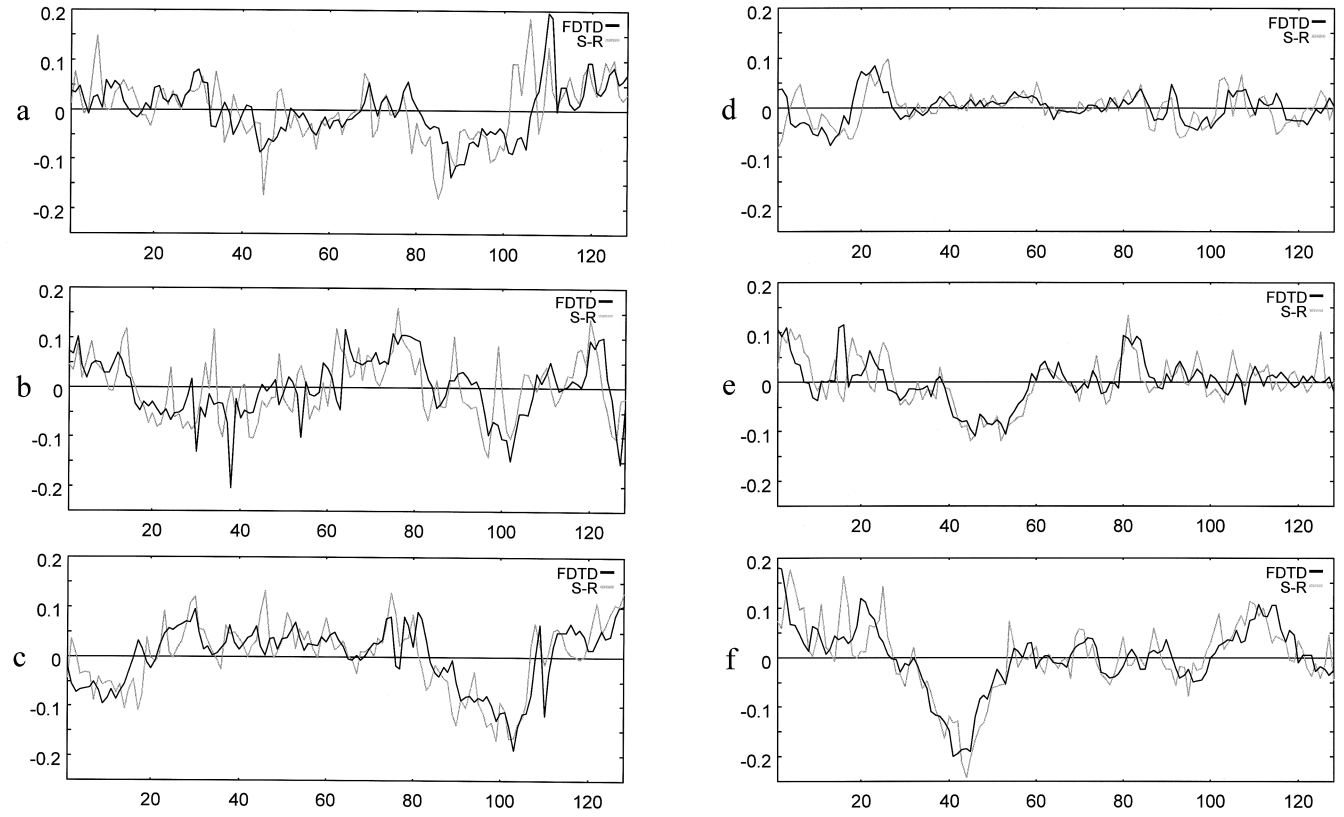


FIG. 7. Arrival time surfaces for FDTD and straight-ray (S-R) simulations. In each plot, the horizontal axis is the element number while the vertical axis is the arrival time fluctuation in microseconds, with positive values representing earlier arrivals. (a) Specimen 75hi. (b) Specimen 77ba. (c) Specimen 87de. (d) Specimen 102gh. (e) Specimen 120de. (f) Specimen 120fe.

ages was employed to gain insight into the development of wavefront distortion. This visualization clearly showed multiple scattering effects; for instance a secondary wavefront caused by scattering from one septum would be further scattered and distorted by other septa later in its propagation path. These effects lead to the complex, random wave field that appears behind the principal wavefront in Figs. 4 and 5.

Received wavefronts from the FDTD simulation are presented in Fig. 6 together with the corresponding experimentally measured wavefronts. The waveforms show qualitatively similar features, including the magnitude of large-scale time-shift and amplitude fluctuations as well as the presence of waveform distortion. However, the locations of these features do not match precisely between computed and experimentally measured wavefronts. Possible reasons for this discrepancy are given in Sec. IV.

Distortion statistics for the wavefronts determined by measurements and the FDTD and straight-ray simulations are given in Table II. The statistics of the FDTD results show qualitative agreement with the statistics of the experimental results. However, it may be noted that the arrival time and wave shape distortion in the simulated results are typically smaller than the measured distortion. Arrival time fluctuations from the FDTD results and from the straight-ray results are presented in Fig. 7. The corresponding energy level fluctuations are presented in Fig. 8. The statistics of each simulation are reported in Table II, while correlation coefficients for the arrival time surfaces and the energy level surfaces which result from the two types of simulation are reported in Table III. Arrival time surfaces calculated using the straight-ray method agree well with those predicted using the FDTD method, as seen by the correlation coefficients, which vary between 0.413 and 0.807. For the aperture size of 128 elements, a correlation coefficient with magnitude greater than 0.2875 is significant to a 99.9% confidence level.[42] This agreement provides evidence that time-shift aberration in the abdominal wall is, in many cases, principally associated with large-scale variations in sound speed. However, the energy level fluctuations calculated by the straight-ray approximation are much smaller in magnitude than those from the FDTD results and the experiments. The two sets of simulated energy level fluctuations are also poorly correlated. These results indicate that differences in the attenuation characteristics of the various tissue types at the center frequency of the transmitted pulse are inherently incapable of explaining the amplitude fluctuation that are observed both in experiments[2,3,5,6,22] and in our full-wave computations. Since the asymptotic straight-ray simulation neglected scattering and ray-bending effects that cause waveform distortion, waveform similarity factors reported in Table II are unity for the straight-ray simulation. More realistic waveform distortion was present in the FDTD simulations, but the amount of waveform distortion in the FDTD results, as quantified by the waveform similarity factor, was somewhat smaller than in the measurements.

The energy level fluctuations reported in Table II can be

TABLE II. Statistics of wavefront distortion from measurements (Exp.), finite-difference simulations (FDTD), and straight-ray simulations (S-R).

Specimen number	Specimen thickness (mm)	Data source	Arrival time fluctuations rms value (ns)	Arrival time fluctuations Correlation length (mm)	Energy level fluctuations rms value (dB)	Energy level fluctuations Correlation length (mm)	Waveform similarity factor
75hi	31–34	Exp.	92.7	4.10	3.85	2.99	0.873
		FDTD	53.0	2.72	3.29	1.25	0.957
		S-R	61.9	2.38	0.50	6.73	1.000
77ba	22–29	Exp.	102.7	3.61	3.98	2.38	0.841
		FDTD	59.9	3.75	4.44	1.17	0.951
		S-R	60.8	2.11	0.46	2.29	1.000
87de	26–30	Exp.	73.7	4.74	3.47	2.75	0.866
		FDTD	60.9	8.69	4.18	1.46	0.948
		S-R	67.8	7.66	0.46	5.23	1.000
102gh	17–21	Exp.	38.7	5.56	3.89	3.22	0.943
		FDTD	28.4	4.48	3.10	1.37	0.986
		S-R	32.3	2.24	0.25	2.57	1.000
120de	25–29	Exp.	59.5	5.76	3.07	2.35	0.958
		FDTD	43.6	2.26	3.28	1.38	0.980
		S-R	48.7	3.71	0.54	10.95	1.000
120fe	28–30	Exp.	73.8	8.66	3.66	3.71	0.914
		FDTD	67.1	4.47	3.41	1.30	0.983
		S-R	73.0	7.99	0.68	10.98	1.000

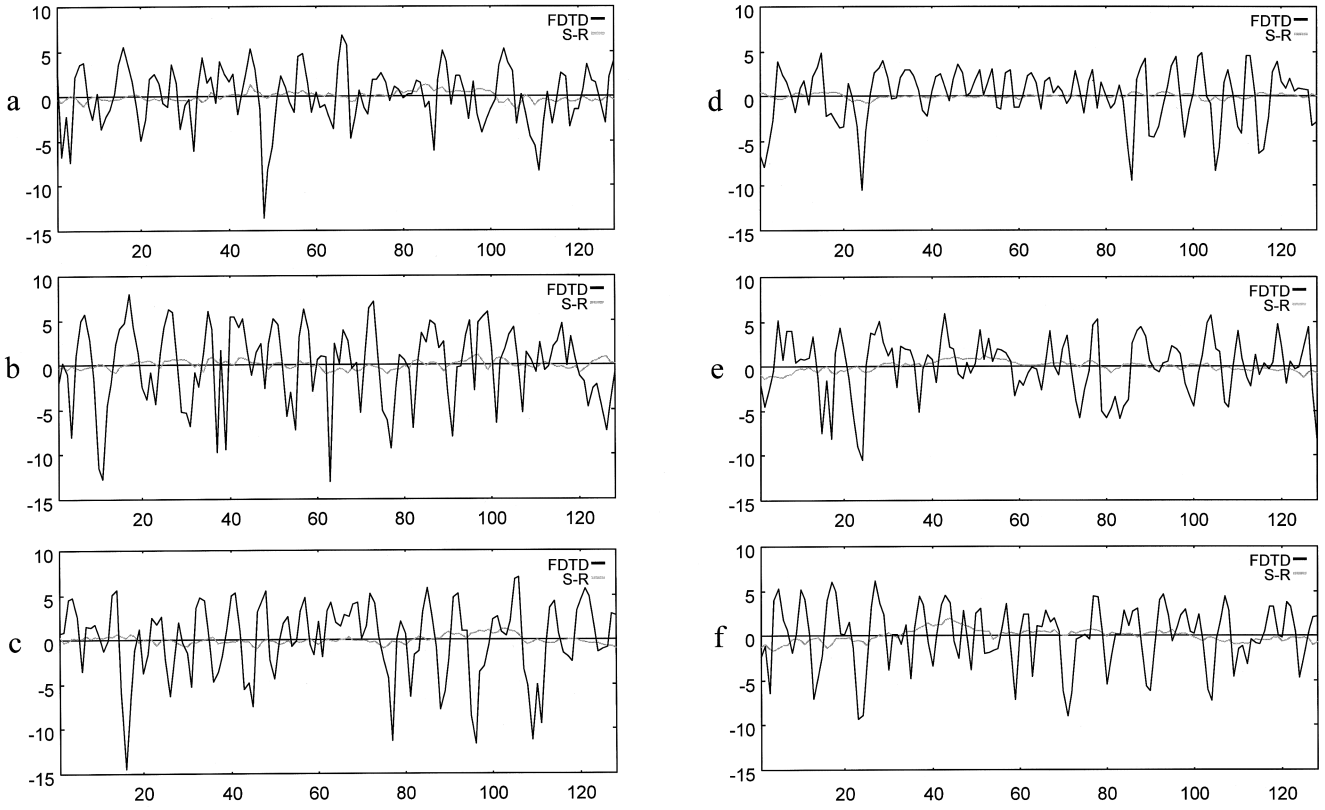


FIG. 8. Energy level surfaces for FDTD and straight-ray (S-R) simulations. In each plot, the horizontal axis is the element number while the vertical axis is the energy level fluctuation in dB, with positive values representing higher amplitudes. (a) Specimen 75hi. (b) Specimen 77ba. (c) Specimen 87de. (d) Specimen 102gh. (e) Specimen 120de. (f) Specimen 120fe.

TABLE III. Correlation coefficients between arrival time surfaces and energy level surfaces from FDTD and straight-ray simulations.

Specimen number	Arrival time correlation	Energy level correlation
75hi	0.508	−0.094
77ba	0.521	−0.051
87de	0.761	0.027
102gh	0.413	0.111
120de	0.693	−0.157
120fe	0.807	−0.065

used to classify ultrasonic scattering in the abdominal wall in terms of common nomenclature for propagation in random media. The rms values of the measured and simulated energy level fluctuations were within the range of 3.1–4.4 dB, which corresponds to a logarithmic amplitude variance of about 0.4–0.5. In a standard reference work on wave propagation in random media, wavefront fluctuations caused by scattering are defined to be weak only if the logarithmic amplitude variance is ''less than about 0.2–0.5.'' Otherwise, strong-fluctuation theory is required.[43] The wavefront distortion effects associated with ultrasonic propagation through the abdominal wall, therefore, should be considered within the range of strong fluctuations that cannot be described fully by weak scattering theory.

IV. DISCUSSION

The simulations and measurements reported here, like the measurements detailed in Refs. 1–6 and 22, model only part of the wavefront distortion that occurs in pulse-echo ultrasonic imaging systems. In the present simulations and experiments, the transmission configuration and broad wavefront model the return path of an initially coherent echo from a scattering site within the abdomen. A transmission configuration is generally used in distortion measurements because this simple arrangement allows the ultrasonic aberration produced by the tissue to be measured directly. However, in pulse-echo imaging, the echo coherence may also be reduced by degradation of the transmit focus. Thus, the distortion measured and calculated here can be regarded as a conservative estimate of the distortion that may be incurred in clinical imaging. In particular, echoes produced by diffuse inhomogeneities illuminated by a poorly-focused beam may be substantially more distorted than those described here.

The measurements and simulations presented here also differ from clinical ultrasonic imaging in that post mortem specimens are employed rather than living subjects. This allows experimental conditions to be controlled precisely and enables tissue morphology to be determined in detail by means of dissection and staining, but implies that the results may not correspond exactly to the clinical situation. It also requires the specimens to be subjected to some form of preservation to prevent degradation during the time between their acquisition, the making of measurements, and imaging for the simulations. The preservation methods used were chosen to minimize their impact on experimental results. For example, the specimens were frozen rather than fixed with formalin before the measurements because formalin fixation alters the elastic properties of tissue. Previous work[2] indicates that freezing affects distortion measurements minimally. The cellular disruption produced by freezing and thawing also should not affect the simulation results because it occurs on scales much smaller than the resolution of the images (0.5- to 20-μm cellular size versus 0.085-mm resolution). The cross sections are finally fixed in formalin because this is required for staining and because it allows them to be kept indefinitely for future reference.

The wavefronts computed using the finite-difference, time-domain method exhibit qualitative agreement with measured distortion for the same specimens. However, the simulated results do not show precise quantitative agreement with the measurements. Several reasons for these differences exist.

First, available computational resources limited the present simulations to a two-dimensional geometry, while real ultrasonic propagation through tissue occurs in three dimensions. Although the computation did not precisely simulate the physical situation of three-dimensional wavefront distortion, the two-dimensional computations are believed to yield a qualitatively accurate description of the development of wavefront distortion as an ultrasonic pulse propagates through abdominal wall. Two-dimensional computations may also be appropriate for characterization of the wavefront distortion that affects clinical imaging systems in which two-dimensional sections of tissue are imaged using a beam focused in the elevation direction. The lack of three-dimensionality is a likely explanation for some of the observed differences between the calculated and measured wavefronts.

Although care was taken to ensure that the sound speed, density, and attenuation parameters used were representative values for the tissues employed, no single group of parameters can characterize all tissue of a given type. As seen in Refs. 21, 37–39, and 41, measured ultrasonic tissue properties vary considerably among individual tissue specimens and different measurement techniques. The tissue model used in this study is also limited by the absence of tissue microstructure below a resolution of about 0.085 mm (corresponding to the 300 d.p.i. resolution of the scanned images employed) and point-to-point variations in sound speed and density within individual tissue types. Although these variations can be quantitatively estimated using acoustic microscopy,[21,39] implementation of this process with high resolution is not practically feasible for large cross sections like those used in the present study. Further limitations arose from the neglect of anisotropy in tissue properties. An extended tissue model could incorporate these effects into similar computations, given that the orientation of each tissue type were known. The approximate nature of the tissue model presented here is an additional cause for the lack of precise quantitative agreement between the simulated and measured wavefronts.

A high degree of uniformity in the received waveforms is necessary for accurate estimation of pulse arrival times when a correlation-based estimation method is used.[44] Table II and Fig. 6 indicate that the measured waveforms are more distorted than those produced by the finite-difference, time-

domain simulation. This may be related to the difference in dimensionality of the distortion encountered in the two cases. Whatever the cause, the result is that the arrival times estimated for the FDTD data are likely to be more accurate than those determined for the corresponding measurements. This could also account for the lower arrival time fluctuation rms values obtained for the FDTD simulation results.

Several other assumptions of the FDTD model were considered not to affect the results significantly. First, no damping was formally incorporated into the finite-difference time-domain model, although the computational method used is slightly dissipative (see Ref. 25 for a discussion of numerical dissipation in finite-difference algorithms). However, the straight-ray attenuation calculations confirm that amplitude fluctuations caused by varying tissue attenuation are much smaller than those caused by scattering effects, so that any artifacts caused by the absence of explicit absorption or the presence of numerical dissipation did not significantly impact the results. The shear elasticity of tissue was neglected in our calculations because tissue has been found to be well approximated by a fluid model at frequencies used in diagnostic ultrasound.[45,46] Therefore, the absence of attenuation and shear elasticity from our FDTD simulations, as well as the small numerical dissipation associated with the finite-difference algorithm, do not affect the validity of our distortion results.

Despite the limitations of the model employed, the presence of realistic time-shift aberrations, amplitude drop outs, and waveform fluctuations indicate that the FDTD method and tissue model employed in this study are sufficient to explain the principal characteristics of wavefront distortion produced by the abdominal wall. The model is, therefore, a useful tool for investigation of the physical causes of wavefront distortion. In particular, the results have shown the formation of large-scale time-shift aberrations with propagation through slowly-varying tissue structures as well as strong amplitude and waveform distortion caused by scattering from inhomogeneities within subcutaneous fat. The computational model presented here is also potentially useful as a tool for the simulation of realistic scattering data for tests of focusing and imaging algorithms.

Since many previous studies of ultrasound-tissue interaction have been based on weak scattering models such as the Born and Rytov approximations,[20,21,47] it is notable that the present calculations show strong scattering and multiple scattering effects not predicted by such simple methods. For example, the large amplitude drop outs seen in the transmitted wavefronts are inconsistent with the Born *ansatz* that the incident wavefront is unperturbed by the scattering medium. Furthermore, examination of the internally scattered fields indicates that a single portion of an ultrasonic wavefront may incur scattering by several inhomogeneities over the large thickness of the abdominal wall, so that single-scattering approximations are seen to provide incomplete representations of the wave effects occurring in real tissue.

Finally, the results reported here clearly show that simple phase-screen and amplitude-screen models do not fully characterize ultrasonic propagation through the abdominal wall. The straight-ray simulation performed in this study is equivalent to a simulation using both a phase and an amplitude screen at the receiving aperture, with time-shift and amplitude-shift values determined by integration of the local tissue sound speed and attenuation along straight propagation paths. Figures 7 and 8 and Tables II and III show that this simple model successfully predicts the magnitudes and large-scale trends of time-shift fluctuations, but is unable to predict energy-level fluctuations and waveform distortions that occur both in experimental measurements and in the full-wave computations reported here. More sophisticated models that include a phase screen some distance from the aperture do predict amplitude and waveform distortion[48–50] and are thus more successful in characterizing propagation through the abdominal wall, as seen from the improved performance of correction algorithms incorporating these models.[6,13] However, no single phase screen, whether alone or in combination with an amplitude screen, can cause cumulative propagation effects and strong scattering effects like those obtained using the present full-wave model and illustrated by Figs. 4 and 5. These effects could be represented, at least in part, by more sophisticated models that include multiple phase screens or numerical ray tracing, but since such models are based on further approximations to the wave propagation equations considered here, they are less complete than the present full-wave approach.

The observation that single phase- and amplitude-screen models are incomplete implies that methods of wavefront distortion correction employing these models should not be expected to restore an ultrasonic focus aberrated by the abdominal wall to its diffraction-limited form. Such algorithms, which were originally developed to correct weak fluctuations such as may occur in radio astronomy, are not fully applicable to correction of strong scattering effects like those shown here to occur within the abdominal wall. Still, the performance of aberration correction algorithms to date suggests that improved models of propagation through tissue, such as the model of a phase screen away from the receiving aperture, have resulted in improved aberration correction. Aberration correction methods that incorporate some of the wave effects shown here could result in further improvements in ultrasonic focusing through the abdominal wall.

V. CONCLUSIONS

A new model for ultrasonic propagation through abdominal wall has been presented. The model is based on realistic cross-sectional tissue maps and on a full-wave solution of the acoustic propagation equations for an inhomogeneous fluid medium.

The model produces wavefront distortion that is statistically similar to experimentally measured wavefront distortion for the same tissue specimens. Some major features of the experimentally measured distortion are reproduced; however, accurate simulation of the precise structure of received wavefronts may require three-dimensional computation as well as more complete tissue models.

Visualization of ultrasonic wavefronts inside the abdominal wall shows the development of time-shift fluctuations as the wavefront propagates through large-scale tissue variations. Examination of the internal pressure fields also

shows that large amplitude and waveform fluctuations can be caused by scattering from high-contrast inhomogeneities such as vertically-aligned septa within the subcutaneous fat layer.

Comparison of the finite-difference, time-domain results with results of a straight-ray approximation suggests that a significant portion of the time-shift aberration caused by the abdominal wall can be explained using simple phase-screen models. However, these models, even those that incorporate tissue attenuation variations, are unable to completely predict amplitude and waveform distortion caused by the abdominal wall. Since these latter distortion effects are associated with strong scattering, complete correction of amplitude and waveform distortion is a challenge that may not be fully attainable using currently available methods of aberration correction that model the aberrating layer as a phase screen alone or in combination with an amplitude screen. Progress toward the goal of diffraction-limited focusing through the abdominal wall may be achieved by incorporation of wave effects such as strong scattering into aberration correction algorithms.

ACKNOWLEDGMENTS

The authors thank Dr. Leon Metlay, Mark Burkard, Cari Kelly, Brian Williams, Georgia Gianakakos, and Kaushik Mittra for assistance in creating accurate tissue maps of the abdominal wall cross sections. Dong-Lai Liu is thanked for providing the original two-dimensional version of the distortion estimation software using the reference waveform method. Funding for this investigation was provided by the University of Rochester Diagnostic Ultrasound Research Laboratory Industrial Associates, NIH Grants Nos. DK 45533 and HL 50855, US Army Grant No. DAMD-17-94-J-4384, and the F. V. Hunt Fellowship of the Acoustical Society of America. Computations were performed at the Cornell National Supercomputing Facility, which is supported in part by the National Science Foundation, New York State, and the IBM Corporation, and at the Center for Academic Computing of the Pennsylvania State University.

[1] Y. Sumino and R. C. Waag, ''Measurements of ultrasonic pulse arrival time differences produced by abdominal wall specimens,'' J. Acoust. Soc. Am. **90**, 2924–2930 (1991).
[2] L. 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).
[3] L. 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).
[4] P. D. Freiburger, D. C. Sullivan, B. H. LeBlanc, S. W. Smith, and G. E. Trahey, ''Two dimensional ultrasonic beam distortion in the breast: In vivo measurements and effects,'' Ultrason. Imaging **14**, 398–414 (1992).
[5] Q. Zhu and B. D. Steinberg, ''Large-transducer measurements of wavefront distortion in the female breast,'' Ultrason. Imaging **14**, 276–299 (1992).
[6] L. M. Hinkelman, D.-L. Liu, R. C. Waag, Q. Zhu, and B. D. Steinberg, ''Measurement and correction of ultrasonic pulse distortion produced by the human breast,'' J. Acoust. Soc. Am. **97**, 1958–1969 (1995).
[7] G. Kossoff, D. A. Carpenter, D. E. Robinson, D. Ostry, and P. L. Ho, ''A sonographic technique to reduce beam distortion by curved interfaces,'' Ultrasound Med. Biol. **15**, 375–382 (1989).
[8] S. W. Smith, G. E. Trahey, and O. T. von Ramm, ''Phased array ultrasound imaging through planar tissue layers,'' Ultrasound Med. Biol. **12**, 229–243 (1986).
[9] M. O'Donnell and S. W. Flax, ''Phase-aberration correction using signals from point reflectors and diffuse scatterers: Measurements,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **35**, 768–774 (1988).
[10] L. Nock, G. E. Trahey, and S. W. Smith, ''Phase aberration correction in medical ultrasound using speckle brightness as a quality factor,'' J. Acoust. Soc. Am. **85**, 1819–1833 (1989).
[11] D. Rachlin, ''Direct estimation of aberrating delays in pulse-echo imaging systems,'' J. Acoust. Soc. Am. **88**, 191–198 (1990).
[12] M. Hirama, O. Ikeda, and T. Sato, ''Adaptive ultrasonic array imaging through an inhomogeneous layer,'' J. Acoust. Soc. Am. **71**, 100–109 (1982).
[13] D.-L. Liu and R. C. Waag, ''Correction of ultrasonic wavefront distortion using backpropagation and a reference waveform method for time-shift compensation,'' J. Acoust. Soc. Am. **96**, 649–660 (1994).
[14] C. Dorme and M. Fink, ''Ultrasonic beam steering through inhomogeneous layers with a time reversal mirror,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **43**, 167–175 (1996).
[15] S. Krishnan, P.-C. Li, and M. O'Donnell, ''Adaptive compensation for phase and magnitude aberrations,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **43**, 44–55 (1996).
[16] D. E. Robinson, L. S. Wilson, and G. Kossoff, ''Shadowing and enhancement in ultrasonic echograms by reflection and refraction,'' J. Clin. Ultrasound **9**, 181–188 (1981).
[17] E. E. Sauerbrei, ''The split image artifact in pelvic ultrasonography: the anatomy and physics,'' J. Ultrasound Med. **4**, 29–34 (1985).
[18] C. W. Manry, Jr., and S. L. Broschat, ''FDTD simulations for ultrasound propagation in a 2-D breast model,'' Ultrason. Imaging **18**, 25–34 (1996)
[19] A. P. Berkhoff, P. M. van den Berg, and J. M. Thijssen, ''Simulation of wave propagation through aberrating layers of biological media,'' 1994 IEEE Ultrason. Symp. Proc. **3**, 1797–1800 (1994).
[20] R. C. Waag, J. O. Nilsson, and J. P. Astheimer, ''Characterization of volume scattering power spectra in isotropic media from power spectra of scattering by planes,'' J. Acoust. Soc. Am. **74**, 1555–1571 (1983).
[21] F. L. Lizzi, M. C. Rorke, D. L. King, E. J. Feleppa, A. Kalisz, J. B. Sokil-Melgar, and J. Nowakowski, ''Simulation studies of ultrasonic backscattering and B-mode images of liver using acoustic microscopy data,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **39**, 212–226 (1992).
[22] B. S. Robinson, A. Shmulewitz, and T. M. Burke, ''Waveform aberrations in an animal model,'' 1994 IEEE Ultrason. Symp. Proc. **3**, 1619–1624 (1994).
[23] A. D. Pierce, *Acoustics: An Introduction to its Physical Principles and Applications* (Acoustical Society of America, Woodbury, New York, 1989), 2nd ed., pp. 14–17.
[24] R. W. MacCormack, *Lecture Notes in Physics* (Springer-Verlag, Berlin, 1971), Vol. 8, p. 151.
[25] D. Gottlieb and A. Turkel, ''Dissipative two-four methods for time-dependent problems,'' Math. Comput. **30**, 703–723 (1976).
[26] V. 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).
[27] P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (McGraw-Hill, New York, 1968), p. 464.
[28] R. L. Burden, J. D. Faires, and A. C. Reynolds, *Numerical Analysis* (Prindle, Weber, and Schmidt, Boston, 1978), p. 380.
[29] The hemispheric source was constructed by Hans Persson of the Department of Electrical Measurements, Lund University, Lund, Sweden.
[30] The linear array employed was a Toshiba model PLE-308M.
[31] L. M. Hinkelman, L. A. Metlay, C. J. Churukian, and R. C. Waag, ''Modified Gomori trichrome stain for macroscopic tissue slices,'' J. Histotech. **19**(4), 321–323 (1996).
[32] The images were processed using Adobe Photoshop, version 3.0.
[33] E. Turkel, ''On the practical use of high-order methods for hyperbolic systems,'' J. Comp. Phys. **35**, 319–340 (1980).
[34] N. Bilaniuk and G. S. K. Wong, ''Speed of sound in pure water as a function of temperature,'' J. Acoust. Soc. Am. **93**, 1609–1612 (1993).
[35] *Handbook of Chemistry and Physics*, edited by R. C. Weast (CRC, Boca Raton, 1985), p. F-10.
[36] K. H. Herzfeld and T. A. Litovitz, *Absorption and Dispersion of Ultrasonic Waves* (Academic, New York, 1959), pp. 353–361.
[37] S. A. Goss, R. L. Johnston, and F. Dunn, ''Comprehensive compilation of

empirical ultrasonic properties of mammalian tissues,'' J. Acoust. Soc. Am. **64**, 423–457 (1978).
[38] S. A. Goss, R. L. Johnston and F. Dunn , ''Compilation of empirical ultrasonic properties of mammalism tissues II,'' J. Acoust. Soc. Am. **68**, 93–108 (1980).
[39] J. 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. Derm. **8**(5), 615–623 (1987).
[40] R. E. Neuman and M. A. Logan, ''The determination of collagen and elastin in tissues,'' J. Biol. Chem. **186**, 549–556 (1950).
[41] H. Q. Woodard and D. R. White, ''The composition of body tissues,'' Br. J. Radiol. **59**(708), 1209–1219 (1986).
[42] P. R. Bevington, *Data Reduction and Error Analysis for the Physical Sciences* (McGraw-Hill, New York, 1969), pp. 310–311.
[43] A. Ishimaru, *Wave Propagation and Scattering in Random Media* (Academic, New York, 1978), Vol. 2, Chap. 20.
[44] G. C. Ng, S. S. Worrell, P. D. Freiburger, and G. E. Trahey, ''A comparative evaluation of several algorithms for phase aberration correction,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **41**, 631–643 (1994).
[45] L. A. Frizzell, E. L. Carstensen, and J. F. Dyro, ''Shear properties of mammalism tissues at low MHz frequencies,'' J. Acoust. Soc. Am. **60**, 1409–1411 (1976).
[46] E. L. Madsen, H. J. Sathoff, and J. A. Zagzebski, ''Ultrasonic shear wave properties of soft tissues and tissuelike materials,'' J. Acoust. Soc. Am. **74**, 1346–1355 (1983).
[47] R. C. Waag, D. Dalecki, and W. A. Smith, ''Estimates of wave front distortion from measurements of scattering by model random media and calf liver,'' J. Acoust. Soc. Am. **85**, 406–415 (1989).
[48] A. Hewish, ''The diffraction of radio waves in passing through a phase-changing ionosphere,'' Proc. R. Soc. London, Ser. A **209**, 81–96 (1951).
[49] R. P. Mercier, ''Diffraction by a screen causing large random phase fluctuations,'' Proc. Camb. Philos. Soc. A **58**, 382–400 (1962).
[50] A. D. Pierce, ''Spikes on sonic boom pressure waveforms,'' J. Acoust. Soc. Am. **46**, 1366–1380 (1968).

The longitudinal vibration response of a curved fiber used for laser ultrasound surgical therapy

S. N. Makarov
Faculty of Mathematics and Mechanics, State St. Petersburg University, Bibliotechnaya, 2, St. Petersburg-Petrodvoretz 198904, Russia

M. Ochmann
Technische Fachhochschule Berlin, Fachbereich Mathematik und Physik, Luxemburger Strasse 10, D-13353 Berlin, Germany

K. Desinger
Laser- und Medizin-Technologie gGmbH, D-12207 Berlin, Germany

(Received 25 July 1996; accepted for publication 13 March 1997)

A theoretical study of the longitudinal vibration response of a bent fiber used as an active element of a medical applicator for Laser Ultrasound Surgical Therapy (LUST) is presented. The fiber cross section ($\varnothing 300-1000\ \mu$m) is much smaller than the longitudinal wavelength ($\lambda = 15-20$ cm). Wave propagation in the bent fiber is described by the governing second-order equations of motion which neglect the flexure effect. The range of the validity of those equations is discussed. In contrast to numerous works on bent rods, the case of an arbitrary continuous curvative distribution along the fiber is investigated. A simple analytical formula for the transfer function, i.e., the ratio of displacements at the working end of the fiber divided by those at the driven end, is obtained. The transfer function depends on frequency, fiber length, output impedance, loss factor, and the mean-square curvative of the fiber. The behavior of this function is investigated applied to some fibers whose lengths are of the order 1 m. The driven frequency is of about 10–50 kHz. If the displacement at the driven end of the fiber is known, the power output of the applicator can be found from the known values of the tissue impedance and the transfer function. © *1997 Acoustical Society of America.* [S0001-4966(97)02607-6]

PACS numbers: 43.80.Gx, 43.80.Sh, 43.40.Cw, 43.20.Mv [FD]

INTRODUCTION

An active element of a promising medical applicator for Laser Ultrasound Surgical Therapy (LUST)[1] is a thin silica glass fiber (300–1000 μm), actually used for laser transmission, which transfers the acoustic energy from a piezoelectric ultrasound transducer to biological tissue. The fiber consists of an amorphous quartz-core and -cladding and a plastic coating to enhance the mechanical stability. For the ultrasound transmission, core and cladding of the fiber can be seen as a homogeneous structure and will therefore be referred to as core. Desinger *et al.*[1] have shown the feasibility of ultrasound and laser transmission via silica glass fibers. This new endoscopic application system would enable surgeons to simultaneously use laser and ultrasound technique for minimal invasive therapy. The ultrasound tissue interaction is based on the well known CUSA technology (Cavitation Ultrasound Surgical Therapy) in medicine which enables the surgeon to cut/desintegrate various types of tissue with different degrees of effectiveness. Typical parameters of the standard CUSA application are a sound frequency range of 17–40 kHz and mechanical elongations between 20 and 350 μm. The coatings of the used fibers are made of polymers (polyimid, acrylate).

For the straight fiber waveguide at low frequencies, the effect of coating is taken into account by an effective value of the bar velocity. The familiar theory of longitudinal acoustic waves in a rod can then be applied[2] to predict the vibratory response of the fiber. The effect of viscoelasticity can further be included.[3] In the case of harmonic excitation, it implies operating with a complex Young's modulus. If the transducer vibration is known and it is not influenced by fiber loading (a massive ultrasound transducer), this vibratory response characterizes the power output of the whole applicator.

An important problem concerns taking into account fiber bending which may appear due to applications in endoscopy. The dynamics of (mainly in-plane) curved rods has been the subject of much research. Comprehensive reference and theory manuals are found in the literature[4–16] related not only to quasi-longitudinal modes but also to quasi-flexural modes.[4–7] For a curved rod, the longitudinal and flexural waves cannot propagate independently of one another[8,9,13] so that only various composite ''longitudinal-flexural'' modes can exist, i.e., the displacements have both a longitudinal and a transversal component. In this paper, we consider only one such mode which degenerates to the longitudinal bar mode at zero curvative and it is therefore called the quasi-longitudinal mode.

The phase (and group) velocity of the quasi-longitudinal mode in a bent fiber depends on its local curvative and on the frequency.[9] If the fiber contains a bend of constant curvative jointed with straight segments without nodes, the phase velocity within the bend remains constant but it is different from bar velocity. The fiber thus resembles a straight rod

with a jump of the phase velocity (material properties). As a result, wave reflection or wave conversion may appear while passing throughout the bend.[14,15] The reflection clearly hinders the energy transfer throughout the fiber. More precisely, the reflection leads to a phase shift between the stress and the displacement at the driven end of the fiber which reduces the acoustic power. The same situation takes place if, instead of one bend, an arbitrary curvative distribution along the fiber is studied.

Therefore, there is some doubt that a curved fiber geometry would be appropriate to achieve an effective energy transfer to the tissue. The main purpose of this paper is to study this problem quantitatively. We consider a thin in-plane curved fiber, whose diameter is much smaller than the wavelength, and ignore the radial inhomogeneity of the acoustic field in the fiber. An approximate theory is used[9] which neglects the effect of flexure. The viscoelasticity is taken into account by a complex value of the Young's modulus. The boundary condition at the point of contact with the tissue is that of the impedance form and only the active impedance is taken into consideration. The effect of the acoustic energy radiation from the fiber walls into the surrounding medium (biological tissue) is ignored.

The main quantity studied in this paper is the longitudinal vibration response of a finite fiber. A formula will be derived for the corresponding transfer function, i.e., the ratio of displacements (or accelerations) at the working end of the fiber divided by those at the driven end. It depends on frequency, fiber length, output impedance, loss factor, and a mean-square curvative of the fiber. The behavior of this function is investigated applied to some fibers whose lengths are of the order of 1 m and the driven frequency is of about 10–50 kHz. The fiber cross section is arbitrary, but the thin fiber condition must always be satisfied. Since the displacement at the driven end of the fiber is known, the power output can then be obtained from the known values of the tissue impedance and the transfer function.

The present study does not include two important effects which are viscous and acoustic energy losses on the walls of a fluid-loaded curved fiber. For quasi-longitudinal low-frequency vibrations, the viscous effect can still be estimated relatively simply, much as the elegant analysis of Nagy and Nayfeh precludes.[17] However, the acoustic energy loss requires consideration of the nonzero transverse displacement [cf. Eq. (10) below] even at small frequencies. The corresponding analysis becomes somewhat similar to that of Dayal[18] and will be the subject of further investigation.

I. OUTLINE OF THE THEORY

Consider at first a straight (cylindrical) fiber of length L at low frequencies, i.e., the fiber whose diameter is much smaller than the wavelength. Acoustic propagation is characterized by three decoupled modes which are the longitudinal (lowest axial-radial) mode, the (lowest) torsional mode, and the (lowest) flexural mode. The fiber properties are given by the Young's modulus E, the density ρ, the cross-sectional area A, and the moment of inertia of the cross section I. Denoting the axial and transverse displacements by ν and w, one has for the longitudinal extension mode and for the flexural mode[2]

$$\rho \frac{\partial^2 \nu}{\partial t^2} - E \frac{\partial^2 \nu}{\partial x^2} = 0, \quad \rho A \frac{\partial^2 w}{\partial t^2} + EI \frac{\partial^4 w}{\partial x^4} = 0, \tag{1}$$

respectively. In contrast to Eqs. (1), the extensional and flexural waves for the curved fiber cannot be exactly decoupled. If the local curvative σ is different from zero but remains constant, the corresponding set of coupled equations takes the form[14–16]

$$\begin{aligned} \rho \frac{\partial^2 \nu}{\partial t^2} - E \frac{\partial}{\partial s}\left(\frac{\partial \nu}{\partial s} + \sigma w\right) &= \frac{EI}{A} \sigma \frac{\partial^2}{\partial s^2}\left(\sigma \nu - \frac{\partial w}{\partial s}\right), \\ \rho \frac{\partial^2 w}{\partial t^2} + E\sigma\left(\frac{\partial \nu}{\partial s} + \sigma w\right) &= \frac{EI}{A} \frac{\partial^3}{\partial s^3}\left(\sigma \nu - \frac{\partial w}{\partial s}\right). \end{aligned} \tag{2}$$

In Eqs. (2), s denotes the centerline arc length so that $\nu = \nu(s,t)$ and $w = w(s,t)$. Setting $\sigma = 0$ one obtains Eqs. (1) from Eqs. (2).

Let us assume that a primary excited wave is an extensional wave. In the bent area, this wave produces reflected extensional and flexural waves as well as transmitted extensional and flexural waves. In accordance with the investigation made in Ref. 14, the effect of flexure for the primary extensional wave becomes relatively small in a wide range of frequency (cf. Figs. 2 and 4 of this reference) if

$$r_0 \sigma < 0.05, \tag{3}$$

where $r_0 = \sqrt{I/A}$ is the radius of inertia. For sharp bends, the inequality (3) is clearly broken. Nevertheless, even in this case, the acoustic energy transferred to flexural waves is considerably smaller in comparison with the energy of extensional waves.[14] For a fiber of circular cross section, $r_0 = 0.5r$ where r is the cross-section radius. Denoting by $R = 1/\sigma$ the radius of curvative one obtains from Eq. (3)

$$\frac{r}{R} < 0.1. \tag{4}$$

The cross-section radius of a fiber r usually does not exceed 1 mm. Therefore, inequalities (3) and (4) are both satisfied if the curvative radius of a bend R is larger than 10 mm, which is obviously true in many cases of practical interest.

Neglecting the flexure items in Eqs. (2) one thus has

$$\begin{aligned} \rho \frac{\partial^2 \nu}{\partial t^2} - E \frac{\partial}{\partial s}\left(\frac{\partial \nu}{\partial s} + \sigma w\right) &= 0, \\ \rho \frac{\partial^2 w}{\partial t^2} + E\sigma\left(\frac{\partial \nu}{\partial s} + \sigma w\right) &= 0. \end{aligned} \tag{5}$$

Equations (5) describe only quasi-longitudinal waves in a bent fiber, which all degenerate to the bar mode at zero curvative σ (transverse displacement w becomes zero). In its present form, they are valid not only for a bend with the constant curvative but also in the case of arbitrary bending as it was shown in Ref. 9 (see also Ref. 19). On the other hand, Eqs. (5) are the exact consequence of the complete equations of motion for bent rods [cf. Ref. 7, Eq. (17)] if one takes a thin rod condition into account and puts the moment of in-

ertia and shearing force resultant being zero.[20] We will use those equations to study the longitudinal vibration response of a finite fiber with arbitrary curvative distribution.

Assuming a harmonic time dependence $\exp(j\omega t)$, designating $\nu=\overline{\nu}(s)\exp(j\omega t)$ and $w=\overline{w}(s)\exp(j\omega t)$, one obtains from Eqs. (5)

$$-\rho\omega^2\overline{\nu}-E\,\frac{d}{ds}\left(\frac{d\overline{\nu}}{ds}+\sigma\overline{w}\right)=0,$$
$$\rho\omega^2\overline{w}-E\sigma\left(\frac{d\overline{\nu}}{ds}+\sigma\overline{w}\right)=0. \tag{6}$$

It follows from the second of Eqs. (6) that

$$\left(\frac{\rho}{E}\,\omega^2-\sigma^2\right)\overline{w}=\sigma\,\frac{d\overline{\nu}}{ds}. \tag{7}$$

This expression for $\overline{w}$ is then substituted into the first of Eqs. (6) which yields

$$-k^2\overline{\nu}-\frac{d}{ds}\left(\frac{d\overline{\nu}}{ds}+\frac{\sigma^2}{k^2-\sigma^2}\,\frac{d\overline{\nu}}{ds}\right)=0,$$

where $k=\omega/c$ with $c=\sqrt{E/\rho}$ is the bar velocity of the straight rod. Hence, an equation for the complex longitudinal displacement $\overline{\nu}(s)$ is obtained in the form

$$\frac{d^2\overline{\nu}}{ds^2}+(k^2-\sigma^2)\overline{\nu}+\frac{2\sigma(d\sigma/ds)}{k^2-\sigma^2}\,\frac{d\overline{\nu}}{ds}=0. \tag{8}$$

If the curvative σ is constant, Eq. (8) is reduced to

$$\frac{d^2\overline{\nu}}{ds^2}+(k^2-\sigma^2)\overline{\nu}=0 \tag{9}$$

and predicts dispersion of the phase velocity $c_p=\omega/\sqrt{k^2-\sigma^2}$ and the group velocity $c_g=c^2\sqrt{k^2-\sigma^2}/\omega$ of the curved fiber with respect to frequency. The phase velocity is always greater than the bar velocity whereas the group velocity is always smaller. Both velocities approach the bar velocity with increasing frequency. There is cutoff at small frequencies which is characterized by the equality of the wavelength of the Young's wave to the radius of curvative.[8,9]

The complex transverse displacement $\overline{w}(s)$ is given by Eq. (7) as follows:

$$\overline{w}=\frac{\sigma}{(k^2-\sigma^2)}\,\frac{d\overline{\nu}}{ds}. \tag{10}$$

It is worth noting that the transverse displacement is expressed in terms of longitudinal strain and vanishes at some spatial point if the local curvative $\sigma(s)$ becomes zero at this point. In accordance with the approach of Ref. 9, the longitudinal stress component T_{ss} can be obtained in the form

$$T_{ss}=E\left(1+\frac{\sigma^2}{(k^2-\sigma^2)}\right)\frac{d\overline{\nu}}{ds}. \tag{11}$$

This relation is simply found from the expression for the normal force resultant [cf. Eq. (3) in Ref. 14] with taking into account Eq. (10). The acoustic power transferred throughout the fiber is given by

$$P=T_{ss}(j\omega\overline{\nu})^*, \tag{12}$$

where the asterisk denotes the complex conjugation. Equation (12) is valid up to second order of approximation. The contribution of the transverse stress into Eq. (12) would be a third-order quantity with respect to the wave amplitude and it is therefore omitted.

II. BOUNDARY CONDITIONS

Since, for the quasi-longitudinal acoustic wave, both longitudinal and transverse displacements are nonzero, boundary conditions might appear more complicated than for a straight rod. However, according to Eq. (10), the transverse displacement is secondary and it is expressed in terms of the corresponding longitudinal quantities. No additional boundary conditions should therefore appear in comparison with the case of a straight rod.

To illuminate this point of view assume that the local curvative of the fiber $\sigma(s)$ and its derivative $d\sigma(s)/ds$ are zero at driven and working ends, i.e.,

$$\sigma(s)=0,\quad \frac{d\sigma(s)}{ds}=0;\quad s=0,L. \tag{13}$$

It means that the fiber is quite straight at the ends. In practice, such a condition usually takes place for a bent fiber in endoscopy applications. If it is not the case, this condition may be achieved for any curved fiber by an appropriate infinitesimally small transform of the fiber position close at the ends. Hence, it is a reasonable assumption in our case, too. The transverse displacement and its derivative are also equal to zero at $s=0,L$ according to Eq. (10). Therefore, boundary conditions involve only the longitudinal quantities.

The boundary condition at the driven end of the fiber reads

$$\overline{\nu}(s=0)=\overline{\nu_0}, \tag{14}$$

where $\overline{\nu_0}$ is the input displacement amplitude (the vibration of the ultrasound transducer). The boundary condition at the working end is that of impedance form

$$T_{ss}=Zj\omega\overline{\nu}, \tag{15}$$

where Z is the longitudinal acoustic impedance of the tissue at the point of contact. By the use of Eq. (11) at $\sigma(L)=0$ one has from Eq. (15)

$$\left.\frac{d\overline{\nu}}{ds}\right|_{s=L}=jk\,\frac{Z}{\rho c}\,\overline{\nu}(s=L), \tag{16}$$

where ρc is the rod impedance. For a composite fiber, the rod impedance may be calculated from the well-known low-frequency theory of clad rods.[21]

The dimensionless output impedance of the fiber is as follows,

$$\overline{Z}=\frac{Z}{\rho c}, \tag{17}$$

and depends both on fiber and medium properties. For glass fibers operating with kidney stones, $\overline{Z}\approx 0.6$ whereas, for soft tissues, $\overline{Z}\approx 0.3-0.1$. The output impedance is taken as a real quantity so that no reactive part changing the resonant frequency of the fiber appears.

In the following, we will always use boundary conditions (14) and (16). The condition of zero termination curvative Eq. (13) applied to fibers with arbitrary bending makes the calculations much more easy (cf. Sec. IV).

III. INCLUSION OF VISCOELASTICITY

The effect of a lossy acoustic material on harmonic wave propagation may be accounted for by introducing the complex Young's modulus E^* instead of E in the form

$$E^* = E(1+j\eta) = E' \exp(j\theta), \tag{18}$$

where $\eta = \tan(\theta)$ is the loss factor and θ is the corresponding phase lag. Equations (8) and (9) can be rewritten in the same form if one replaces k by k^* where $k^* = \omega/c^*$ and $c^* = \sqrt{E^*/\rho}$. Now k^* is the complex propagation constant and c^* is the complex bar velocity. The effective complex bar velocity of a cladded fiber is given by (cf. Ref. 21)

$$c^* = \sqrt{\frac{E_2(1+j\eta_2)[((r_2/r_1)^2-1)+E_1(1+j\eta_1)/(E_2(1+j\eta_2))]}{\rho_2[((r_2/r_1)^2-1)+\rho_1/\rho_2]}}, \tag{19}$$

where the indexes 1 and 2 are related to core and coating, respectively, and $r_{1,2}$ are the corresponding cross-section radii. The effective loss factor is as follows $\eta_{\text{eff}} = 2\,\text{Im}[c^*]/\text{Re}[c^*]$, where Re and Im denote the real or imaginary part, respectively.

IV. TRANSFER FUNCTION

Much the same as Jimeno-Fernandez *et al.*[22] we introduce the transfer function T which is the ratio of displacements (or accelerations) at the working end of the fiber divided by those at the driven end. The transfer function T is given by

$$T = \frac{\overline{\nu(s=L)}}{\overline{\nu_0}}. \tag{20}$$

The transfer function actually characterizes the resonance response of the fiber with respect to particle displacement. For the straight rod, the transfer function depends on the frequency, the rod length, the output impedance, and the loss factor. In our case, an additional dependence on the curvative should appear. By analogy with Eq. (20), one might also consider the resonance response in terms of acoustic strain. The power at the working end of the fiber is found from Eqs. (12) and (15):

$$P = Z\omega^2 \overline{\nu\nu^*}.$$

The use of Eqs. (17) and (20) yields

$$P = \rho c\, \omega^2 \overline{Z} |\overline{\nu_0}|^2 |T|^2. \tag{21}$$

It follows from Eq. (21) that the power is characterized by the transfer function if the transducer vibration and the output impedance are known.

A. Solution at constant curvative

We are looking for a solution of Eq. (9) in the form

$$\overline{\nu} = \overline{\nu_0}\{C_1 \exp(-jKs) + C_2 \exp(jKs)\}, \tag{22}$$

where $K = \sqrt{(k^2-\sigma^2)}$. The boundary conditions (14) and (16) give the coupled equations with respect to C_1 and C_2

$$\begin{aligned} &C_2 = 1 - C_1, \\ &-jK \exp(-jKL)C_1 + jK \exp(jKL)C_2 \\ &\quad = jk \frac{Z}{\rho c} (\exp(-jKL)C_1 + \exp(jKL)C_2), \end{aligned} \tag{23}$$

which yield

$$C_1 = \left(1 + \frac{1+k\overline{Z}/K}{1-k\overline{Z}/K} \exp(-2jKL)\right)^{-1}, \quad C_2 = 1 - C_1. \tag{24}$$

Here, $\overline{Z}$ is given by Eq. (17). The transfer function from Eq. (20) then becomes

$$T = \frac{F \exp(-jKL)}{1 + F \exp(-2jKL)}, \quad F = \frac{2K}{K - k\overline{Z}},$$

or, in a simplified form,

$$T = \frac{1}{\cos(KL) - (jk\overline{Z}/K)\sin(KL)}. \tag{25}$$

The transfer function of the straight waveguide for the purely propagating mode ($\overline{Z}=1$) is given by $T = \exp(jkL)$. Hence, its absolute value equals unity and only the phase shift with respect to the driven end becomes apparent. In general, the function T shows a rather difficult behavior.

If the loss factor η is small, one obtains from Eq. (18) $c^* = c(1+j\eta/2)$ and $k^* = k(1-j\eta/2)$. The dimensionless quantities are introduced

$$\xi = kL, \quad \phi = L\sigma, \tag{26}$$

characterizing the effect of the frequency and the curvative, respectively. Performing the calculations above with k^* instead of k the transfer function takes the form

$$\begin{aligned} T_\eta(\xi, \phi, \overline{Z}) &= \frac{1}{\cos\kappa - (j\xi\overline{Z}/\kappa)\sin\kappa}, \\ \kappa &= \sqrt{(\xi^2(1-j\eta) - \phi^2)}. \end{aligned} \tag{27}$$

Equation (27) can be evaluated numerically. However, before doing it, we consider the case of a fiber with an arbitrary

curvative distribution. In this case, it will also be possible to get an analytical solution for the transfer function.

B. Solution at arbitrary continuous curvative

Equation (8) may be rewritten in the form

$$\frac{dy}{ds}=-L(k^2-\sigma^2)x-\frac{2\sigma(d\sigma/ds)}{k^2-\sigma^2}\,y,\quad \frac{dx}{ds}=\frac{1}{L}\,y,$$

where $y=L(d\overline{\nu}/ds)$ and $x=\overline{\nu}$. Denote by $\mathbf{D}=(x,y)$ the solution consisting of x and y. One has

$$\frac{d\mathbf{D}}{ds}=\begin{bmatrix}0 & 1/L\\ -L(k^2-\sigma^2) & -\dfrac{2\sigma(d\sigma/ds)}{k^2-\sigma^2}\end{bmatrix}\mathbf{D}.$$

It follows that

$$\mathbf{D}=\hat{A}(s)\mathbf{D}_0 \tag{28}$$

with

$$\hat{A}(s)=\exp\left(\int_0^s\begin{bmatrix}0 & 1/L\\ -L(k^2-\sigma^2) & -\dfrac{2\sigma(d\sigma/ds)}{k^2-\sigma^2}\end{bmatrix}ds\right),$$

$$\mathbf{D}_0=(d_1,d_2),\quad 0\leqslant s\leqslant L. \tag{29}$$

Here, d_1 and d_2 are two arbitrary constants. To find the transfer function, we are only interested in the solution at two points $s=0$, L. From Eq. (28), one gets

$$\mathbf{D}(s=0)=\mathbf{D}_0,$$

$$\mathbf{D}(s=L)=\exp\begin{bmatrix}0 & 1\\ -L^2\left(k^2-\dfrac{1}{L}\displaystyle\int_0^L\sigma^2\,ds\right) & \ln(k^2-\sigma^2)|_0^L\end{bmatrix}\mathbf{D}_0.$$

In accordance with Eq. (13), the last diagonal term of the matrix in square brackets vanishes so that one has

$$\begin{aligned}&\mathbf{D}(s=0)=\mathbf{D}_0,\\ &\mathbf{D}(s=L)=\exp\begin{bmatrix}0 & 1\\ -L^2\left(k^2-\dfrac{1}{L}\displaystyle\int_0^L\sigma^2\,ds\right) & 0\end{bmatrix}\mathbf{D}_0.\end{aligned} \tag{30}$$

Equation (30) allows us to calculate the transfer function if we use the standard eigenvalue technique to evaluate the matrix exponent. However, the result immediately becomes clear if one writes the solution for the circle of constant curvative in the form of Eq. (30), too. Since the curvative is the same at both ends of the circle, the last diagonal term vanishes again. The only distinction thus is found in the substitution

$$\frac{1}{L}\int_0^L\sigma^2\,ds\rightarrow\sigma^2, \tag{31}$$

which is made upon the exponent in Eq. (30). Therefore, the transfer function looks like Eq. (27), but instead of the constant dimensionless curvative $\phi=L\sigma$ the averaged dimensionless curvative Φ should be introduced as follows:

$$\Phi=\sqrt{L\int_0^L\sigma^2\,ds}. \tag{32}$$

The losses are again taken into account by using k^* instead of k. Equation (27) is thus transformed to

$$T_\eta(\xi,\Phi,\overline{Z})=\frac{1}{\cos\,\kappa-(j\xi\overline{Z}/\kappa)\sin\,\kappa},$$

$$\kappa=\sqrt{(\xi^2(1-j\,\eta)-\Phi^2)}. \tag{33}$$

Equations (32) and (33) are the main results of this study. They show that the transfer function of a curved fiber only depends on the mean-square curvative of the fiber. Hence the specific bending geometry (the number of circles, their sequence, and their radii) plays no role if the total mean-square curvative remains unchanged. This circumstance allows us to obtain some general results for curved fibers discussed in Secs. V–VII below.

The present solution is convincing if

(i) $\sigma(s)<k$ for any internal integration point s.
(ii) $\sigma(s)$ is a continuous function along the fiber.

Otherwise, the integrand in the last diagonal term of matrix A given by Eq. (29) will be a singular function at some $s\in[0,L]$. The integration requires a special analysis of each internal peculiar point. The calculations are no longer straightforward and the approach used just above may be incorrect. Below in this section, we consider some examples where conditions (i) and (ii) are not satisfied.

C. The effect of curvative jumps

Let us assume that $\sigma(s)$ has a jump at a certain internal point s_0 but remains continuous at other points [condition (ii) is broken]. Such a behavior would correspond to a circular segment ideally jointed with a straight segment. Then, instead of the equality

$$-\int_0^L\frac{2\sigma(d\sigma/ds)}{k^2-\sigma^2}\,ds=\ln(k^2-\sigma^2)|_0^L=0, \tag{34}$$

one has

$$-\int_0^L\frac{2\sigma(d\sigma/ds)}{k^2-\sigma^2}\,ds=-\ln(k^2-\sigma^2)|_{s_0-0}^{s_0+0}+\ln(k^2-\sigma^2)|_0^L\neq 0, \tag{35}$$

where the symbols $s_0\mp 0$ indicate that the values of the corresponding function are taken just before the jump and just after the jump, respectively. Herewith, Eqs. (30) and (33) are not valid anymore.

It is, however, believed that for flexible bent fibers the continuity condition is usually achieved in practice and we do not need to take curvative jumps into consideration. The jumps are only necessary while studying reflection and transmission on the basis of a simplified theory which operates with a number of segments of constant curvatives.[9,14,15]

D. The effect of local cutoff

Now we check condition (i) above. Note that the equality $\sigma = k$ determines the cutoff frequency of the fiber of constant curvative. Hence, for a fiber with variable curvative, condition (i) may be called a local cutoff condition. If it is not satisfied, the integral

$$-\int_0^L \frac{2\sigma(d\sigma/ds)}{k^2-\sigma^2}\,ds \tag{36}$$

can only exist as a generalized function of parameter k. Consider a symmetrical curvative distribution with respect to the central point $s = L/2$ of the form

$$\sigma^2 = k^2(1-(2s/L-1)^2), \tag{37}$$

where

$$\sigma(0) = \sigma(L) = 0, \quad \sigma(L/2) = k. \tag{38}$$

One writes

$$-\int_0^L \frac{2\sigma\, d\sigma/ds}{k^2-\sigma^2}\,ds = 2\int_0^L \frac{1}{s-L/2}\,ds = 0 \tag{39}$$

in the generalized sense of Cauchy's principal value. Unfortunately, we cannot state such equality for an arbitrary curvative distribution. Inclusion of complex viscoelasticity or other perturbation effects (the effect of flexure) may help to avoid generalized functions in the integral (36). A more detailed treatment of this problem is beyond the scope of the present paper.

V. ANALYSIS

A. General

A first estimate, which ranges the domain of the validity of Eq. (33), follows from Eqs. (3) and (4) and states that local curvative radii cannot be very small. Further, we should take into account the local cutoff condition discussed in the previous section. In terms of dimensionless variables from Eq. (26), one has

$$\phi_{\max} < \xi, \tag{40}$$

where $\phi_{\max}$ is the largest curvative along the fiber. This inequality gives another limitation to the use of Eq. (33) which is usually stronger than Eqs. (3) and (4).

If the fiber curvative is constant, Eq. (33) predicts nearly zero values of the transfer function below the cutoff frequency $\Phi = \xi$ (cf. next subsection). If, however, the bent area is fixed, but the whole fiber length increases due to increase of straight areas, the transfer function (33) tends to the transfer function of a straight fiber to within a change in phase. Such a property becomes clear if one writes

$$\Phi^2 = O(L), \quad \xi^2 = O(L^2), \quad \kappa \to \xi\sqrt{(1-j\eta)}, \quad L \to \infty, \tag{41}$$

which means that the curvative Φ loses its influence on the parameter κ. Herewith, we come to a conclusion which states that a local bend inside of a long straight fiber only weakly affects its transfer function. At first sight, this conclusion might appear incorrect if the driven frequency approximates the cutoff frequency of the bend. To make things clear one should remember the following. Equation (33) is correct under the assumption that the local cutoff condition can never be achieved [cf. Eq. (40)]. Transport of acoustic energy throughout the bend is therefore always available. While the total fiber length increases a number of intermediate resonances appears. This circumstance makes the transport of acoustic energy through the bend more efficient than in the case of a short fiber. At long last, the transfer function can then approximate the corresponding dependence for the straight rod.

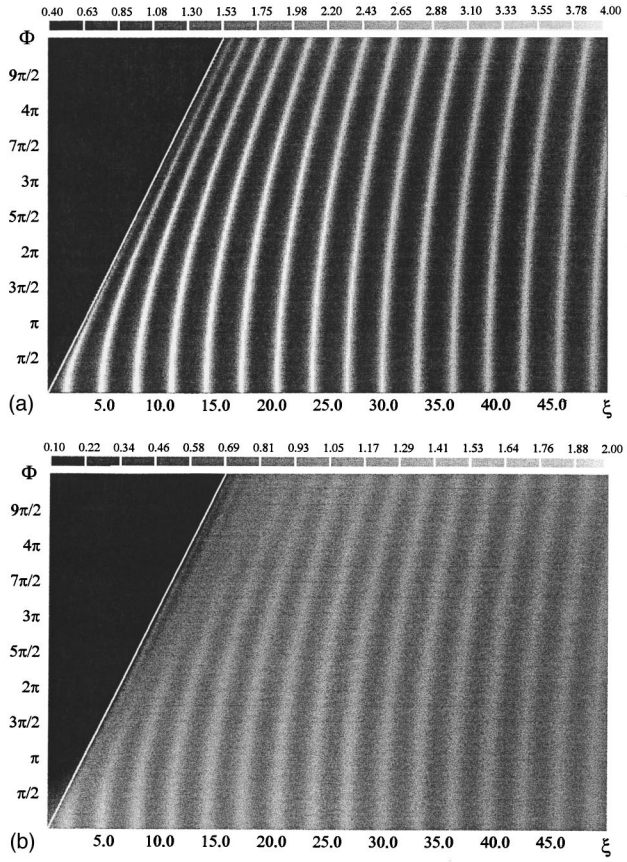


FIG. 1. Transfer function in the plane of the dimensionless fiber length (frequency) ξ and the dimensionless mean-square curvative Φ. The straight white line shows the dependence $\Phi = \xi$. The gray scale gives the transfer function magnitude: (a) $\overline{Z} = 0.1$, scaling is fixed on the interval from 0.4 to 4; (b) $\overline{Z} = 0.6$, scaling is fixed on the interval from 0.1 to 2.

B. Resonances

Figure 1(a) shows a 2-D representation of the absolute value of the transfer function from Eq. (33) at the dimensionless output impedance $\overline{Z} = 0.1$ (soft tissues). The loss factor η is equal to 0.01. The x axis corresponds to the dimensionless fiber length or frequency ξ whereas the y axis describes the dimensionless mean-square curvative Φ. The domain of change in ξ covers glasslike fibers whose lengths are of the order of 1 m and the driven frequency does not exceed 50 kHz. Shading scaling is fixed on the interval from 0.4 to 4 (the values smaller or equal than 0.4 are black whereas the values greater or equal than 4 are white). The straight line outlines the dependence $\Phi = \xi$. Above this line (the upper left corner of the figure), the transfer function is close to zero

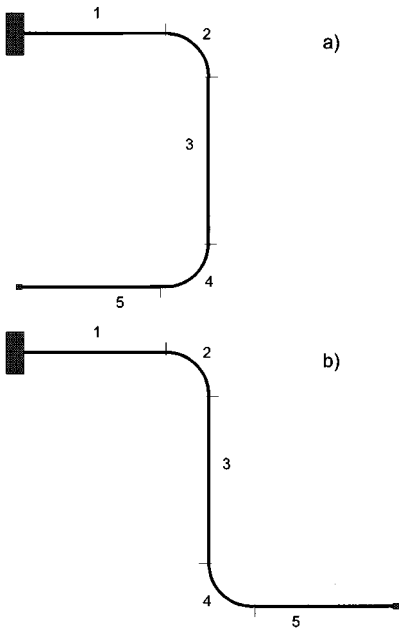


FIG. 2. The geometry of a bent fiber consisting of five different segments 1–5. Two bending combinations are possible. They are shown as (a) and (b), respectively.

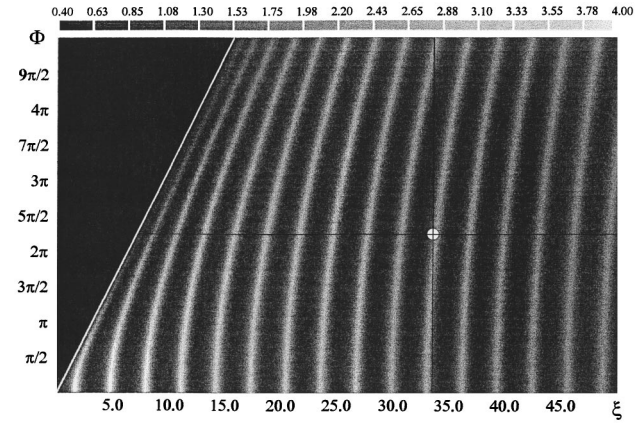


FIG. 3. Transfer function at $\overline{Z}=0.2$ in the plane of the dimensionless fiber length (frequency) ξ and the dimensionless mean-square curvative Φ. The gray scale gives the transfer function magnitude. The scaling is fixed on the interval from 0.4 to 4. The straight white line shows the dependence $\Phi=\xi$. The position of the fiber is indicated by a white circle.

irrespective of the values of the other fiber parameters. The equality $\Phi=\xi$ would therefore be identified as the cutoff of a bent fiber.

To the right of the line $\Phi=\xi$, a number of resonances are seen (they are given by white strips). The resonances give values of the transfer function up to 5. The resonant stripes change their position while the averaged curvative is changed. Hence, if both the driven frequency and the fiber length are fixed, the transfer function also changes its value in dependence on the mean-square curvative of the fiber. In other words, the output efficiency may be different for current fiber geometries due to the change in resonant frequencies. The magnitude of this difference is the ratio of transfer function values at resonances divided by those in the ''dark'' domains. This ratio was found to be up to 4 in Fig. 1(a). Note that the effect of fiber bending on the behavior of resonant stripes is smaller at larger frequencies.

If the output impedance Z is large, the resonances are not quite developed and the effect of bending is probably of lesser importance for the transfer function. The corresponding example is given in Fig. 1(b) which provides the same information as Fig. 1(a) but Z is now equal to 0.6 (kidney stones). The loss factor η is again 0.01.

As a further example, let us consider a fiber of 1 m length at $Z=0.2$. The driven frequency equals 30 kHz. The fiber consists of five different segments with lengths $L_1,\ldots,L_5$ and curvatives $\sigma_1,\ldots,\sigma_5$, respectively. The first, third, and fifth segments have zero curvative (straight lines) whereas the second and fourth segments are the circular segments (Fig. 2). Figure 3 shows a position of this fiber in the plane ξ, Φ and the corresponding transfer function. The circular segment apertures are both equal and are 90°. Similar to Fig. 1(a), shading scaling is fixed on the interval from 0.4 to 4. The other fiber data read $L_1=25$ cm, $L_2=10$ cm, $L_3=30$ cm, $L_4=10$ cm, and $L_5=25$ cm so that both curvative radii are equal to 6.4 cm [still below the limiting value c/ω given by Eq. (40)]. The fiber mean-square curvative is close to a resonance and the transfer function equals 2.27.

VI. COMPARISON WITH EXPERIMENT

There are a number of experiments on wave propagation in waveguides of constant curvative or helical springs (cf., for instance, Ref. 23). Some experiments have been also done to study the effect of variable curvative distribution on the behavior of displacement amplitude at the working end of a fiber or a rod.[24,25] In Ref. 24, the pulse propagation has been investigated along a $\frac{3}{8}$-in. brass-alloy cylindrical rod consisting of an arc and some straight segment. As a theoretical model in the frequency domain, the authors have also applied Eq. (8) to within a change in variables. Excellent agreements between theory and experiment were found. The general result (33) was, however, still unknown, and the numerical solution obtained by synthesis in intervals and successive integration looks very complicated.

Below, we will treat another experiment which studies continuous radiation.[25] The transducer includes piezoceramics which generate the ultrasound and an amplitude transformer. Thin hollow fibers 28.0–79.5 cm in length are made of chrome–nickel–steel alloy with 18% chrome and 10% nickel. The inner diameter of the fiber varies from 0.4 to 1.6 mm whereas the outer diameter changes from 0.8 to 2.4 mm, respectively. The fiber geometry implies a circular segment of constant curvative and one or two straight segments. The bending was simulated by putting a part of the fiber into a PTFE hose which was then stuck to massive circles of various cross-sectional radii. The displacement amplitude at the working end of the fiber was measured as a function of the segment curvative by an electrodynamic sensor. Its sensitivity was about 0.1 mV per 1 μm. The aperture of the circular segment was fixed by 45°, 90°, and 180°, respectively. The working frequency was 17.5 kHz and the bar velocity of the fiber was 4783 m/s. Since the working frequency is reasonably small, the most significant parameter for the theory is

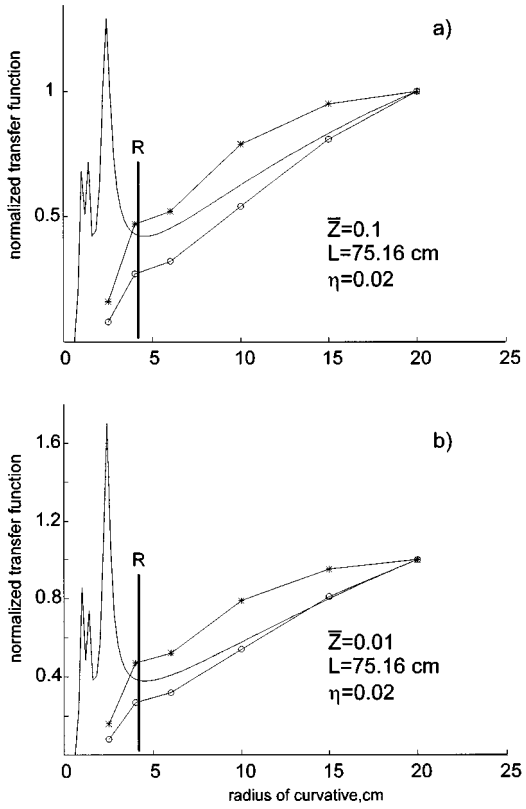


FIG. 4. Transfer function in dependence on segment curvative. The aperture angle of the curved segment is 180°. Asterisks and circles indicate experimental data[25] for the fibers with outer diameters of 0.8 and 1.6 mm, respectively. (a) $\bar{Z}=0.1$ and (b) $\bar{Z}=0.01$. The continuous curve shows the numerical solution. The thick vertical line (R) indicates the local cutoff condition.

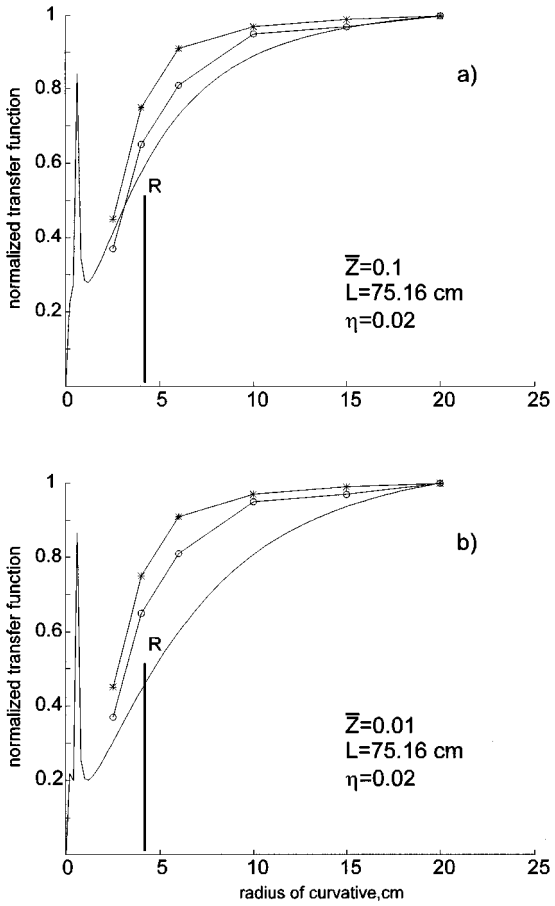


FIG. 5. Transfer function in dependence on the segment curvative. The aperture angle of the curved segment is 45°. Asterisks and circles indicate experimental data[25] for the fibers with outer diameter of 0.8 mm and 1.6 mm, respectively: (a) $\bar{Z}=0.1$ and (b) $\bar{Z}=0.01$. The continuous curve shows the numerical solution. The thick vertical line (R) indicates the local cutoff condition.

the bar velocity of the fiber. The hollow structure of the fiber and its diameter are of less importance and are therefore not involved by Eq. (33).

Before starting with the comparison of theoretical and experimental[25] data, certain remarks should be made concerning the experimental conditions. While performing the experiment in air, the output dimensionless impedance of the fiber should be small enough with respect to unity. Actually, its value depends on the characteristics of the electrodynamic sensor used to measure the displacement[25] and remains unknown. Therefore, we have to check various $\bar{Z}\ll 1$.

Although the total fiber length was not reported explicitly, it was only possible to achieve a curvative radius of 20 cm at the 79.5-cm length cited below in Ref. 25 if the segment aperture equals 180°. However, the nearest resonance of the straight fiber lies at 75.16 cm. A distinction between those values indicates either coupling with the ultrasound transducer or an error in the bar velocity. To get the resonance condition, in the theory we should put $L=75.16$ cm exactly. If we choose the fiber length in such a manner, then the effects of the ultrasound transducer, inaccurate phase velocity, etc., may be ignored to a certain extent.

Figures 4 and 5 show the transfer function normalized to its value at zero curvative in dependence on the segment curvative at 180° and 45° apertures, respectively. Stars and circles indicate experimental data obtained for two fibers of different cross sections. The numerical solution of Eq. (33) is represented by a continuous line. Figures 4(a) and 5(a) correspond to $\bar{Z}=0.1$, whereas Figs. 4(b) and 5(b) correspond to $\bar{Z}=0.01$. The loss factor is 0.02 everywhere. Thick vertical lines show the local cutoff. If the curvative radius is smaller than this value, then Eq. (40) is not satisfied.

It is seen that the theory follows the experimental tendency up to very small radii of curvative for both aperture angles. In the domain of large curvative radii, the dependence on $\bar{Z}$ is rather weak, similar to the dependence on η. The main inaccuracy appears at small radii of curvative where the transfer function and the displacement show unexpected large jumps. The jumps have no physical meaning and can hardly be seen experimentally. They, however, appear only below the local cutoff frequency. In this domain, the present theory can be incorrect since Eq. (40) is not satisfied.

Two important problems are to be discussed at the end of this section. First, it has been found that the transfer function is very sensitive to the resonance condition outlined above. If the resonance condition is slightly broken (for instance, if the fiber length is chosen as 79.5 cm), then the calculated curves diverge from the experimental data. On the other hand, the behavior of the transfer function for various

fiber resonances appears rather similar, especially if higher resonances are dealt with.

The second question concerns the experimentally observed dependence of the transfer function on the position of the circular segment inside the straight fiber (Fig. 17 in Ref. 25). Our model does not predict such a dependence. It should, however, be emphasized that the dependence found implies small radii of curvature (3.5 cm at 17.5 kHz; $c = 4783$ m/s), i.e., again the domain below the local cutoff frequency.

VII. CONCLUDING REMARKS

(i) Once the oscillation of the ultrasound transducer is known, the longitudinal vibratory response of a bent fiber is given by Eq. (33). The vibratory response only depends on the mean-square curvative of the fiber but does not depend on the specific bending geometry under study. This conclusion implies that local bending is not very strong, i.e., the local cutoff condition [Eq. (40)] should be satisfied. Otherwise, a more general theory is necessary, including the flexure effect.

(ii) The vibratory response determines the acoustic power transferred throughout the fiber to some point in the tissue [Eq. (21)].

(iii) Above the local cutoff frequency, the resonant frequencies of the fiber are changed by bending. As a result, considerable decrease in the transfer function may appear in comparison with the case of a straight rod. This effect is of little importance in the situations where the dimensionless output impedance is of the order of unity since the resonances themselves are not quite developed.

Two ways are probably possible to overcome the decrease in the transfer function under presence of fiber bending. The former implies ''tuning up'' the driven frequency in accordance with the current mean-square curvative of the fiber. The latter consists in the use of larger frequencies and higher resonances of the fiber. As Figs. 1 and 3 show, those resonances are somewhat less sensitive to fiber bending than the leading ones.

[1] K. Desinger, J. Tschepe, T. Stein, and G. Müller, ''Laser Ultrasonic Surgical Therapy (LUST),'' in *Minimal Invasive Medizin* [Ecomed Verlagsgesellschaft **7**(3) (1996)]; and in *IEEE 1995 Ultrasonics Symposium Proceedings*, edited by M. Levy and B. R. McAvoy (IEEE, New York, 1995), pp. 1585–1588. See also ''Transmission of ultrasound at high and low frequencies via optical silica glass fibers,'' in *Ultrasonics World Congress 1995 Proceedings*, edited by J. Herbertz (GEFAU, Duisburg, 1995), pp. 1083–1085.

[2] M. C. Junger and D. Feit, *Sound, Structures, and Their Interaction* (MIT Edward Brothers, Cambridge, 1972).

[3] H. Kolsky, *Stress Waves in Solids* (Clarendon, Oxford, 1953), Chap. 5.

[4] W. B. Bickford and B. T. Strom, ''Vibration of plane curved beams,'' J. Sound Vib. **39**, 135–146 (1975).

[5] Y. Tene, M. Epstein, and B. T. Sheinman, ''Dynamics of curved beams involving shear deformation,'' Int. J. Solids Struct. **11**, 827–840 (1975).

[6] Š. Markuš and T. Nánási, ''Vibration of curved beams,'' Shock Vib. Dig. **13**, 3–14 (1981).

[7] J. P. Charpie and C. B. Burroughs, ''An analytical model for the free in-plane vibration of beams of variable curvative and depth,'' J. Acoust. Soc. Am. **94**, 866–879 (1993).

[8] L. Filipczynski, ''Propagation of ultrasonic waves in spirals,'' Proc. Vib. Prob. **3**, 241–251 (1962).

[9] M. A. Isakovich and L. N. Komarova, ''Longitudinal-flexural modes in a slender rod,'' Sov. Phys. Acoust. **14**, 491–494 (1968).

[10] A. E. Green and N. Laws, ''A general theory of rods,'' Proc. R. Soc. London, Ser. A **293**, 144–155 (1966).

[11] N.-C. Huang, ''Theories of elastic slender curved rods,'' J. Appl. Math. Phys. (ZAMP) **24**, 1–19 (1973).

[12] G. Eason, ''Wave propagation in a naturally curved elastic rod,'' J. Sound Vib. **36**, 491–511 (1974).

[13] A. S. Gendy and A. F. Saleeb, ''Vibration analysis of coupled extensional/flexural/torsional modes of curved beams with arbitrary thin-walled sections,'' J. Sound Vib. **174**, 261–274 (1994).

[14] C. M. Wu and B. Lundberg, ''Reflection and transmission of the energy of harmonic elastic waves in a bent bar,'' J. Sound Vib. **190**, 645–659 (1996).

[15] C. M. Wu and B. Lundberg, ''Reflection and transmission of the energy of transient elastic extensional waves in a bent bar,'' J. Sound Vib. **191**, 261–272 (1996).

[16] K. F. Graff, *Wave Motion in Elastic Solids* (Clarendon Oxford, 1975).

[17] P. B. Nagy and A. H. Nayfeh, ''Viscosity-induced attenuation of longitudinal guided waves in fluid-loaded rods,'' J. Acoust. Soc. Am. **100**, 1501–1508 (1996).

[18] V. Dayal, ''Longitudinal waves in homogeneous anisotropic cylindrical bars immersed in fluid,'' J. Acoust. Soc. Am. **93**, 1249–1255 (1993).

[19] The corresponding result of Ref. 9, which coincides with Eqs. (5), is given by equations on the top of p. 492. The sign of the transverse displacement w should be changed in accordance with the commonly used notations. Note that, in this reference, the equations of motion were also presented for arbitrary bending, taking into account the flexural impedance of the rod. However, if the local curvative is put at constant, those equations do not coincide with Eqs. (2) used in this paper. Certain care is therefore required while applying the conclusions of Ref. 9 concerning flexural terms. A comprehensive discussion of the general equations of motion for curved rods was presented in Ref. 12.

[20] The thin rod condition means here that the dimensionless parameter Z in Eq. (11) of Ref. 7 approximates zero. In Eq. (12) of Ref. 7, the sign of the right-hand side should be positive as it follows from the previous consideration [cf. Eq. (10) of Ref. 7]. The coincidence with Eqs. (5) used in the present paper then takes place.

[21] R. N. Thurston, ''Elastic waves in rods and clad rods,'' J. Acoust. Soc. Am. **64**, 1–37 (1978).

[22] I. Jimeno-Fernandez, H. Überall, W. M. Madigosky, and R. B. Fiorito, ''Resonance decomposition for the vibratory response of a viscoelastic rod,'' J. Acoust. Soc. Am. **91**, 2030–2033 (1992).

[23] W. G. B. Britton and G. O. Langley, ''Stress pulse dispersion in curved mechanical waveguides,'' J. Sound Vib. **7**, 417–430 (1968).

[24] D. Y. Hsieh and J. P. Lee, ''Experimental study of pulse propagation in curved elastic rods,'' J. Acoust. Soc. Am. **54**, 1052–1055 (1973).

[25] U. Stumpff, ''Generation and transmission of high-intensity ultrasound in flexible waveguides in the 20 kHz-domain for therapeutic applications,'' Ph.D. thesis, Aachen, Reinisch-Westfälische Technische Hochschule, 1978, 137 pp. (in German).

Bioacoustic behavior of African fishes (Mormyridae): Potential cues for species and individual recognition in *Pollimyrus*

John D. Crawford,[a] Aaron P. Cook,[b] and Andrea S. Heberlein[c]
Department of Psychology and Institute of Neurological Sciences, University of Pennsylvania, 3815 Walnut Street, Philadelphia, Pennsylvania 19104

(Received 2 October 1996; accepted for publication 13 April 1997)

An analysis of the natural bioacoustic signals made by two closely related African fishes (*P. adspersus and P. isidori*) revealed that these species separated along several acoustic dimensions that are likely to be important for species isolation. Both species produced *grunts* that were composed of a trains of pulses, but the pulse repetition rates were distinctly different (56±3 s.d. vs 44±4 s.d. pps). Complex tone bursts (*moans*) were also used, but the species differed substantially in the location of the fundamental peak (240 Hz±12 s.d. vs 332 Hz±34 s.d.). Some *P. adspersus* males sustained these tones for over a second (812 ms±495 s.d.), whereas *P. isidori* produced shorter tones (121 ms±35 s.d.). During interactions with females, the two species produced the grunts and moans in distinct species-typical patterns: *P. adspersus* males alternated grunts with moans and *P. isidori* produced a single grunt followed by a succession of moans. A detailed analysis of identified individual *P. adspersus* showed that acoustic features constituted individual *signatures* which could be used by conspecifics to identify individuals. Grunt spectral peak frequency was shown to be a good predictor of male mass, with peak frequency decreasing at 72 Hz per gram. Simulated *standardized* courtship encounters with females revealed that males differ markedly in their apparent ability to produce sustained moans and it is suggested that this may be particularly important to females in mate selection. [S0001-4966(97)07407-9]

PACS numbers: 43.80.Ka, 43.80.Jz [FD]

INTRODUCTION

The production of acoustic signals for communication is ubiquitous among animals, and the analysis of these signals is fundamental to comparative and evolutionary research on sound communication. Although most readily observed in terrestrial species, acoustic signal production underwater is well known in vertebrate and invertebrate animals, inhabiting both marine and freshwater environments. Acoustic communication in fishes is common (reviews by Fine *et al.*, 1977; Myrberg, 1997). The fishes are of particular interest in the neuroethological analysis of vertebrate acoustic communication due to the relative simplicity of both the signals and the auditory systems of these vertebrates. Sound is transduced at inner ear organs (sacculus, utriculus, or lagena) that are relatively undifferentiated compared to the cochleas of mammals and auditory organs of other vertebrates. Nevertheless, the mechanosensory hair cells which mediate hearing in fish are anatomically and biophysically quite similar to those within the cochleas of mammals and other amniotes; recent studies have shown that the sensory epithelia within the inner ear organs of fishes are heterogenious with at least two distinct classes of hair cells (e.g., Popper *et al.*, 1993; Saidel *et al.*, 1995). After transduction at the ear, auditory information is processed along central nervous system pathways that are essentially similar to the mammalian pathways (McCormick, 1992).

In some fishes there are specialized gas-filled structures coupled to the hair cells of the ear which allow fish to detect pressure (e.g., Fay *et al.*, 1982; Coombs and Popper, 1982; Popper and Fay, 1984), the same acoustic parameter used for sound detection in humans and many other terrestrial animals. Among the fishes with specializations for pressure detection are the African Mormyridae in which the largest structure within each ear is a gas-filled bubble attached to the sacculus (Stipetić, 1939; Fig. 1A in Crawford, 1993). In addition to this specialization of the ear, it has recently been discovered that the mormyrids produce some of the most elaborate acoustic communication signals known in fishes. It has been suggested that these sounds function in mate choice and species isolation (reviewed in Crawford, 1997).

In this paper, we present data which provide further support for these hypotheses. An analysis of the acoustic signals of two closely related mormyrids from West Africa, *Pollimyrus adspersus* and *P. isidori*, shows clear species differences in acoustic signals. The acoustic behavior of breeding *P. adspersus* has been studied in our laboratory for several years. Recent field expeditions to Mali have made it possible to record the acoustic signals of its close relative *P. isidori*. *P. isidori* were studied during their reproductive season in the flood plains of the Central Delta of the Niger River, near Timbuktu. These field recordings of *P. isidori* have, for the first time, allowed comparisons of the acoustic behavior of two closely related species, both of which live in the Niger

[a]Electronic mail: jud@psych.upenn.edu
[b]Current address: Department of Psychology, University of Maryland, College Park, MD 20742-4411.
[c]Current address: Department of Psychology, Stanford University, Stanford, CA 94305-2310.

River and may be sympatric in part of their natural geographic range (Bigorne, 1990).

We first present a comparison of key features of the acoustic signals of the two species. The acoustic separation between the species is clear and supports the hypothesis that the sounds may function in mate choice. An acoustic analysis of identified individual *P. adspersus* is then presented. These analyses show that within this species sounds could be used as individual signatures, and that the characteristics of an individual's sounds may yield clues about its physical condition (mass). Portions of this work have appeared in abstract form (Crawford, 1995; Crawford *et al.,* 1995).

I. MATERIALS AND METHODS

P. adspersus were obtained from the Niger River Basin through importers in Nigeria. Their sounds were recorded in the laboratory during artificial breeding seasons as described in previous papers (e.g., Crawford, 1991) and summarized below. *P. isidori* were studied during a natural breeding season in the field. The conditions under which the sounds were recorded were different, but we have focused our species comparisons on acoustic features which we suspect are particularly salient for the animals's auditory system and which appear relatively independent of physical-acoustic differences in the recording environments. For example, we have confirmed that the temporal patterns within the sounds of *P. adspersus* are essential identical whether a given individual produces these sounds in a large (600-liter) or small (200-liter) tank. Experiments at the field site in Mali have confirmed that *P. isidori* sounds were essentially similar whether recorded in the flood plain or in a small tank, but did show some anticipated changes in spectral content due to environmental filtering (Crawford *et al.*, 1997). The details of waveform fine structure and spectral shape are expected to differ with environment, and were not used in the species comparison presented in this paper. Environmental influences were minimized by recording close to the source as described below, and by recording laboratory sounds at the same water temperature as the field sounds (28 °C). Observations of *P. isidori* that we transported from the field site to our laboratory in Philadelphia have confirmed that acoustic behavior is not changed by extensive air travel; males court females with the same acoustic display observed in Africa.

The species comparison is followed by an examination of the features of sounds which might be used for intraspecific mate choice and individual recognition. This analysis depends on having data from known individuals and is therefore based on our laboratory recordings of identified male *P. adspersus*. In the laboratory, we were able to use playbacks of female electric signals to simulate the stimulus that naturally elicits male sound production (see Crawford, 1991). This allows us to observe sound production under standardized conditions, and to examine male sound production when the duration of the fictive female visit was extended for an unusually long time to *probe* male acoustic performance (details below).

A. Field studies of *P. isidori*

Field recordings of *P. isidori* were made in Mare du Bare, Mali, during the flooding season in August 1994. A complete account of the breeding behavior, ecology, and methodology has been presented in a separate paper (Crawford *et al.*, 1997), and only details essential to the acoustic analyses discussed in the present paper are provided here. Recordings were made at night in a shallow flooded zone fed by waters from the Mayo Ninga River, a tributary of the Niger River. The fresh water was about 2 m deep, 28 °C, and isothermal (±1%) from top to bottom.

Male fish establish territories within the subsurface portions of floating rafts of grass. Hydrophones (Sippican VLAD; sensitivity −40 dB *re*: 1 V per Pa) were suspended 50 cm below the water surface within male territories. We only used sounds recorded with good signal-to-noise ratio in these analyses, and we estimate that they were recorded at distances ranging from 5 to 50 cm from the sound-producing fish; the fish could not be observed visually in this environment. Sounds were either recorded directly to tape or transmitted by radio to a receiver on land and then recorded (see Crawford *et al.*, 1997). Sounds were amplified with a BMA 202 amplifier before being recorded on a Marantz DAT or a Sony WMD6C analog cassette recorder.

B. Laboratory studies of *P. adspersus*

The reproductive behavior of *P. adspersus* has been studied in several laboratories and methods for breeding have been detailed previously (e.g., Crawford, 1992; Kirschbaum, 1987; Bratton and Kramer, 1989). *P. adspersus* were imported from Nigeria (West Africa), and were kept 2–5 adults per large all-glass aquarium (200–625 liters). Reproductive behavior was induced by simulating a rainy season: periodic rain and declining water conductivity. Aquaria were kept at 28 °C (±2%), and conductivity was held at 35 μS/cm. Fish were maintained on a 12-h light–dark cycle.

Acoustic recordings were made as described above for field recordings. A hydrophone was suspended in the center of a male's territory and sounds were recorded at a distance of 5–15 cm from the fish. In addition to recording spontaneous sound production, we used playback methodology (see Crawford, 1992) to electrically simulate the presence of a female on the male's territory. In these weakly electric fish, the female's electric signal is a potent *releaser* for the male's acoustic courtship display. When ready to spawn, females visit the territories of males, producing electric discharges, and males respond with sound production. By using playbacks, we could elicit natural displays to a *standardized female* model, and control where and then the male would produce his sounds; males typically hovered around the container housing the dipole electrode during a playback. The hydrophone was suspended directly above the container. All recording was done during the first few hours of the dark cycle when the fish are most active and naturally engage in courtship (Crawford *et al.*, 1986).

C. Electric stimulation

Previous research has shown that male sound production is released by the higher regular pattern of electric organ discharge (EOD) that females produce during natural courtship interactions; the timing of the EODs is a critical stimulus dimension for the male, and the details of the female's EOD waveform are relatively unimportant in *P. adspersus* (Crawford, 1991). Artificial square or sinusoidal pulses can be substituted for the natural EOD without diminishing the potency of the playback stimulus so long as a uniform repetition rate of about 20 electric pulses per second is used. Based on these findings, we used a stimulator (Grass model S88) to trigger a function generator (Exact model 7056) at 50-ms intervals for trials of 10–60 s, as specified in the results section. The function generator produced a single period sinusoid (200 μs). This signal was passed through a stimulus isolation unit and applied to a dipole electrode (AgCl) to create a naturalistic electric image of a fish (5 mV/cm peak–peak at 10 cm from the dipole). In a few cases we used natural EODs with inter-EOD-interval distributions recorded from females during courtship. The responses to these two types of electric stimuli were the same.

D. Acoustic analyses

Sounds were digitized from tapes using the analog-to-digital converter (A/D) of a Macintosh computer (Quadra 840 AV and Power Macintosh 8100), and sound analysis software developed by the Cornell Bioacoustics Center (*Canary*). Sounds were analog filtered on input to the analog-to-digital converter (100-Hz to 10-kHz pass band) and amplified as necessary to make optimal use of the A/D range. Sounds were usually digitized at 24 kHz with 16-bit sampling. In a few cases 22.45 kHz and 8 bits were used but this did not influence the acoustic measures used here, confirmed by measurements on several sounds digitized in both ways.

Pulse repetition rates within grunts were computed as the reciprocal of the mean inter-pulse-interval. The time from the peak of the first pulse to the peak of the last pulse (=grunt duration) was divided by the number of intervals (= n pulses -1) to get the mean interval (Fig. 2A). Only pulses with at least a 6-dB signal-to-noise ratio were counted. Peak level of the background noise was taken from a 100-ms sample preceding the sound. To examine the interval pattern within individual grunts, the duration of each interval was measured to the nearest 45 μs and the time of the first pulse of each pair was used as the time of occurrence of the interval.

Amplitude spectra were particularly useful for characterizing moans since they were essentially complex tone bursts with distinct peaks in their spectra. The locations of the first and second spectral peaks were measured to the nearest 1.4 Hz, and the bandwidth of the first spectral peak was also measured using a 20-dB criterion (BW_{20}; Fig. 2B). The BW_{20} was defined by the separation (in Hz) of the two points where the spectrum first dropped by 20 dB relative to the peak. Spectra were based on 16 384-point fast Fourier transforms with a Hamming window function applied.

E. Size measurements

The masses of some of the live *P. adspersus* were measured in order to examine relationships with acoustic variables. Mass was recorded to the nearest 0.2 mg with an electronic balance (Denver Instrument Co., XE Series, model 100A).

F. Temperature effects on sounds

Aquarium water temperatures were varied in order to determine how acoustic parameters of sounds might be influenced by body temperature in *P. adspersus*. Since the fish are small (≈8 g) and exothermic, it was assumed that body temperature was equivalent to the temperature of the water near the fish. Water temperature was varied within the 26–29 °C range with the aid of thermostatically controlled water heaters (300 W). Water temperatures were changed gradually so that a new stable temperature was reached within a minimum of 24 h. Temperatures were measured to the nearest 0.1 °C with a digital temperature probe (Orion model 122). The sounds of each fish were recorded at 2–4 different temperatures, according to a unique randomly selected temperature sequence; the number of temperature points varied since some fish failed to produce sounds at the *extreme* temperatures. We computed the slope of the best-fit line relating rate (or fundamental frequency) to temperature for each fish as well as standard physiological Q_{10}'s; the Q_{10} assumes that rate changes exponentially with temperature $Q_{10} = (R_1/R_2)^{(10/t_2 - t_1)}$, where R is rate and t is temperature (Schmidt-Nielsen, 1979).

G. Statistical analysis

Statistical analyses were performed on the Macintosh with Systat v5.2, Statview II and Microsoft Excel. For some

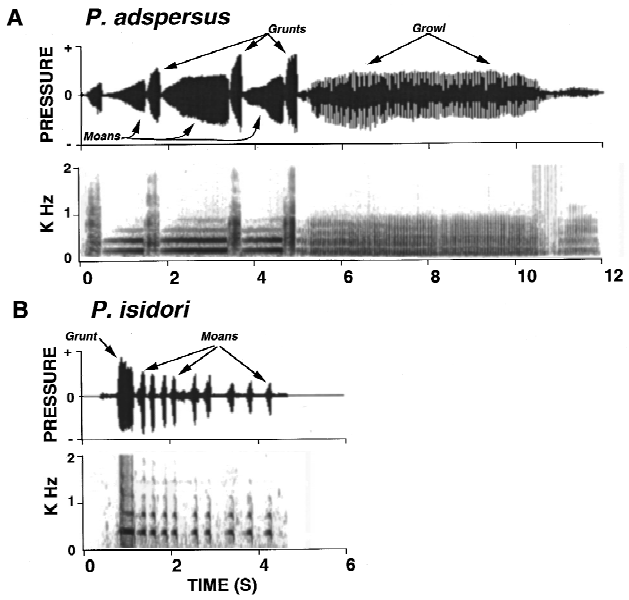


FIG. 1. Acoustic displays of (A) *P. adspersus* and (B) *P. isidori*. In each case, the waveform of the signal is shown on top and the corresponding sonagram is shown below. Sounds were digitized with a 44-kHz sampling rate; the sonagrams are based on 2048-point FFTs, with a frame length of 1024 points, 87.5% overlap, and a Hamming Window.

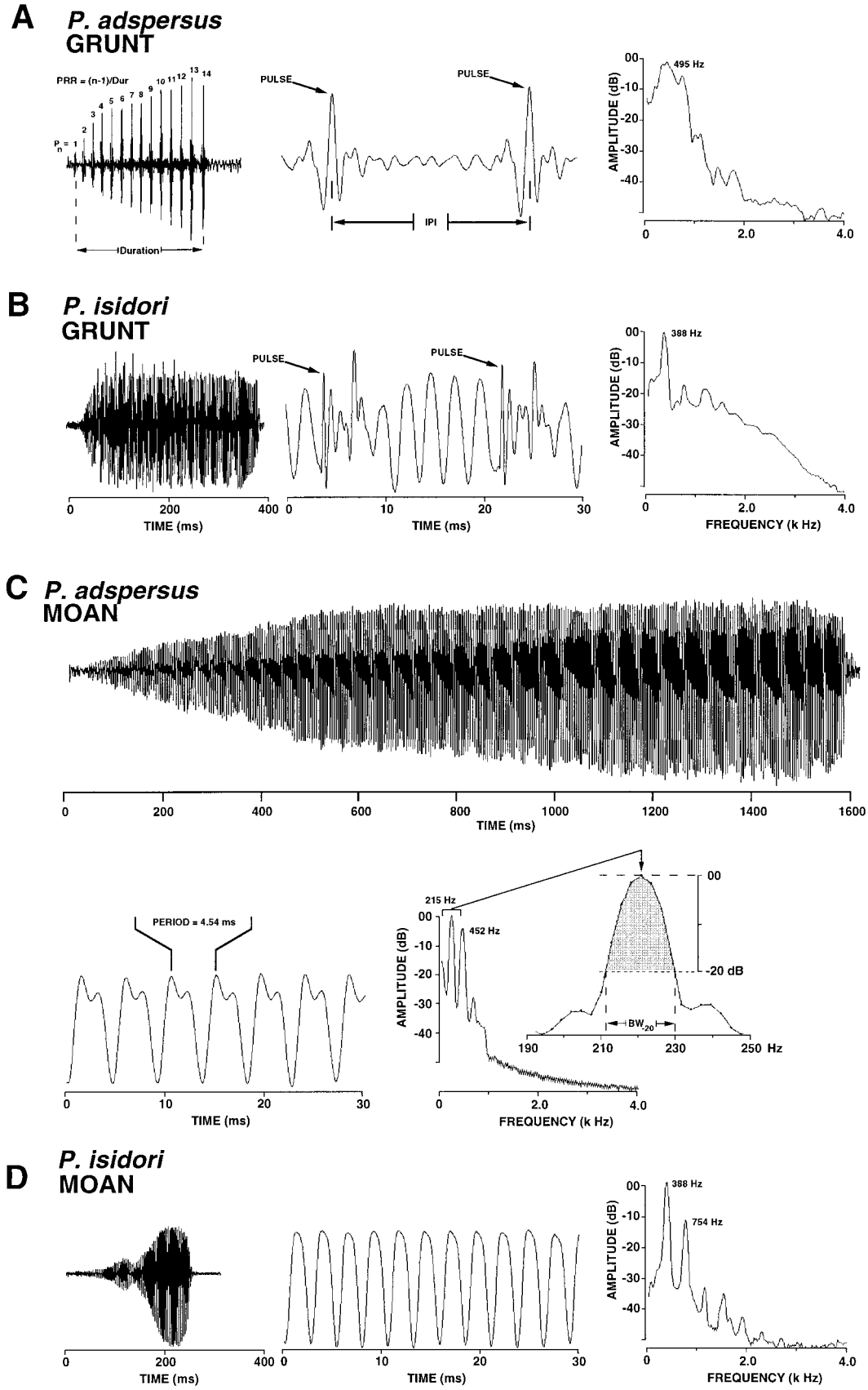


FIG. 2. Characteristics of *grunts* (A, B) and *moans* (C, D) for two species of *Pollimyrus*. In each case, the signal is provided on two different time scales to show the whole signal and the fine structure of the signal. Amplitude spectra are shown at the far right for each sound (2048-point FFT, 1024 point Frame, 87.5% overlap). Measurements of duration and inter-pulse-interval (IPI) are indicated in A, and BW_{20} in C (bottom right). Locations of spectral peaks are provided on the amplitude spectra (e.g., 495 Hz in A). Waveforms were digitally filtered between 100 Hz and 5 kHz.

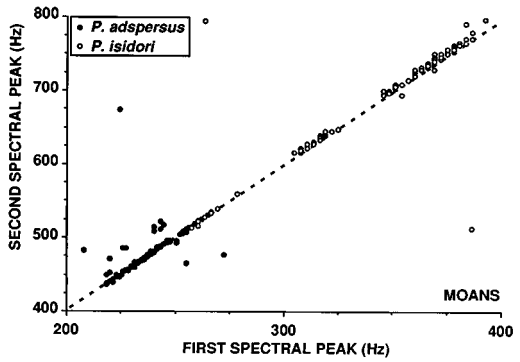


FIG. 3. The first two peaks in the amplitude spectra of moans are related by an integer multiple of two in both species. The dashed line indicates the predicted 2:1 ratio for harmonics, and each point represents the first and second spectral peak for a single moan. One hundred and five moans, 6–8 per male, for each of 15 individual *P. adspersus* are shown as solid points. One hundred and forty moans, from 27 field recording sessions, 1–8 moans per session, at 7 distinct locations, are shown as open circles for *P. isidori*.

of the species comparisons, distributions were not compared statistically because the number of degrees of freedom was not certain (see Fig. 4).

II. RESULTS

A. Species differences

Males of *P. adspersus* and *P. isidori* had distinct species-specific acoustic repertoires (Fig. 1). These species were similar in so far as males of both species composed their acoustic display from one element that was essentially a pulse train with a relatively broad amplitude spectrum (*Grunt*; Fig. 2A, B), and another that was essentially a complex tone burst with two major peaks in the spectrum (*Moan*; Fig. 2C, D; Fig. 3). However, beyond these similarities there were a variety of ways in which the displays clearly differed. The sonic courtship display of *P. adspersus* consisted of a regular alternation between the grunt and the moan during an interaction with a female (Fig. 1A; see also Crawford *et al.*, 1986 and Crawford, 1991). In contrast, *P. isidori* typically produced a single grunt followed by a series of moans (Fig. 1B). The moans of *P. adspersus* were relatively long, often a second or more in duration, whereas the moans of *P. isidori* were typically only a few hundred milliseconds in duration (Fig. 4D). In many cases, the *P. isidori* grunt included a substantial *moanlike* component, visible between successive pulses and preceding the first pulse of this sound, and this yielded a pronounced peak in the spectrum (e.g., Fig. 2B). This was not nearly as apparent in the laboratory recordings of *P. adspersus* grunts (e.g., Fig. 2A), although the grunts of some individuals did show a less pronounced spectral peak at the fundamental of the moan. In *P. adspersus*, the grunt–moan sequence was often followed by an epoch of pulsatile *growling*. Growling was not produced in association with the grunt–moan display in *P. isidori* (see also Crawford *et al.*, 1997).

P. adspersus and *P. isidori* used different pulse repetition rates (PPR) in their grunts. The mean PRR for *P. adspersus* was 12 pulses per s (pps) higher than that of *P. isidori* (56 pps±3 s.d. vs 44 pps±4 s.d.) and there was little overlap in their distributions (Fig. 4A). These two species also produced moans with distinctly different spectral peak frequencies. The mean frequency for the first spectral peak (Fig. 2C) was lower by 92 Hz in *P. adspersus* than in *P. isidori* (240 Hz±12 s.d. vs 332 Hz±34 s.d.), and again there was reasonably clean separation of the two distributions (Fig. 4B). These two species produce grunts of similar duration (Fig. 4C) but the mean duration of the *P. adspersus* moan was about 7-fold greater than that of *P. isidori* (Fig. 4D). In summary, male *P. adspersus* produced fast grunts and long, slow moans compared to their congener *P. isidori*.

B. Individual differences in *P. adspersus*

The acoustic signals of individual males provided cues sufficient for individual recognition. For example, individuals differed significantly in the pulse repetition rates used in grunts (ANOVA; $p \ll 0.001$; $F = 7.78$; df=16 males-1; 8 grunts per individual), and in the first spectral peak of moans (ANOVA; $p < 0.001$; $F = 3.15$; df=16 males-1; 8 moans per individual). When individuals were plotted according to these acoustic dimensions, the separation of individuals was clear in most cases (Fig. 5). Field recordings of *P. isidori* have suggested that under natural conditions a given individual can detect, at most, 4–5 neighbors at one time (Crawford *et al.*, 1997), thus potentially making the task of individual recognition even simpler than indicated by the plot of 10 fish shown here. Males also differed significantly in the spectral bandwidth (BW_{20}) of moans (ANOVA; $p \ll 0.001$; $F = 4.8$; df=16 males-1; 8 moans per individual), inter-grunt-interval (ANOVA; $p < 0.001$; $F = 3.41$; df=16 males-1; 8 intervals per individual), and grunt duration (ANOVA; $p < 0.001$; $F = 4.25$; df=16 males-1; 8 grunts per individual). Thus males differ along a variety of acoustic dimensions that might be useful to other fish for individual recognition. In the laboratory, or in the field (*P. isidori*), human listeners can usually easily discriminate 2–4 individuals even in the absence of binaural cues (pers obs).

Our analyses of grunts revealed that, although males differed significantly in both pulse repetition rate and grunt duration, individuals tended to exhibit less variation in repetition rate than in duration. A given individual may produce grunts varying in duration by 100 ms or more, but this is achieved primarily by adding intervals (or subtracting), and not by slowing down the repetition rate (Fig. 6). For each individual, duration was strongly correlated with the number of intervals used (e.g., Fig. 6A). The mean slope was 15.4 ms/interval ±2.15 SE, and the sample of slopes was significantly different from zero ($p = 0.002$; $t = 7.2$, df=5 males-1). Although interval duration did increase steadily over the course of the grunt (Fig. 6B, C), this effect accounted for only a few ms ($\approx$150 μs per interval).

Acoustic features of the grunt may also provide information about the size of the sound producing individual. When the peak frequency of the amplitude spectrum of the grunt (mean of 8 grunts/male; 9 males) was regressed on body mass, a negative relationship was found ($R^2 = 0.52$; $p < 0.03$; $F = 7.53$; df=1,7). Peak frequency declined at about 72 Hz per gram of body mass (Fig. 7A). A similar analysis of variance was performed for moan peak frequency, but was

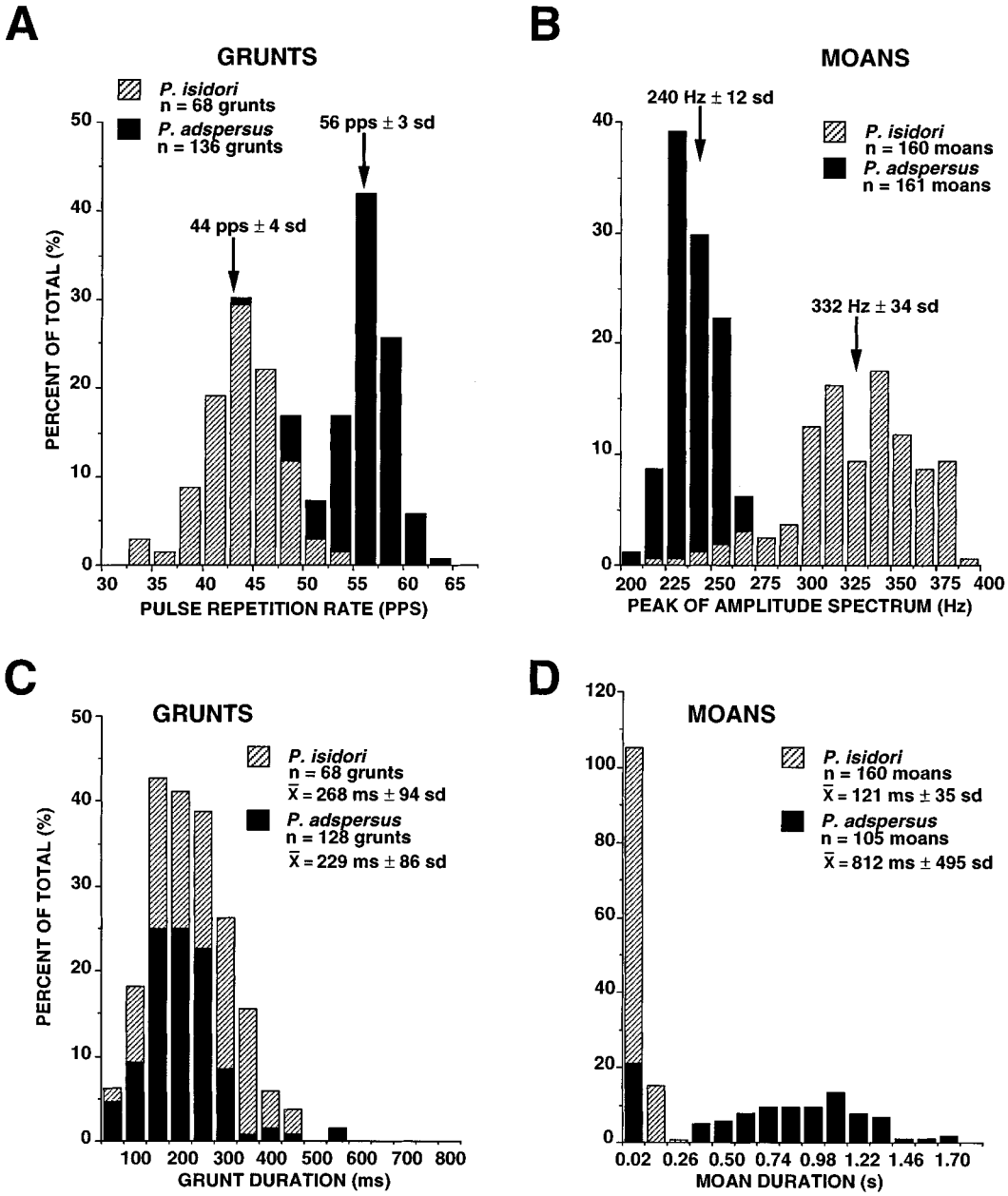


FIG. 4. Parameters of grunts (A, C) and moans (B, D) for *P. isidori* and *P. adspersus*. Summary statistics are provided for each frequency distribution. Bars are stacked. The sum of bar heights is 100% for each species and the corresponding number of sounds (n) is provided in each case. Note that the *P. adspersus* sounds are from a known number of identified individuals, whereas the *P. isidori* sounds were recorded in the field from an uncertain number of free animals.

not significant ($R^2=0.19$; $p<.24$; $F=1.66$; df=1,7; Fig. 7B). Thus grunt peak frequency might provide information about male size, with larger males producing deeper grunts.

C. Acoustic responses to sustained electric stimuli

During natural interactions with females, male *P. adspersus* usually produced a continuous sequence of grunts and moans (see Fig. 1, 0–5 s). These courtship interactions on the male's territory often lasted for 10–20 s, and when the female left the territory the male immediately stopped the grunt–moan display, and often began a protracted bout of growling (Fig. 1, 5–11 s). Males differed in their ability to concentrate acoustic energy at a fundamental frequency during the production of moans in the inter-grunt period. We postulated that this might be an aspect of acoustic behavior that would be significant to potential mates and carried out experiments to explore differences between males in acoustic responses. We investigated male performance during simulated female visits. We were particularly interested in the male's sound production over the course of what might be an unusually long visit by a female. Since it is known that the female's electric signal elicits the male's sonic courtship behavior (Crawford, 1991), we used electric playback methods to probe males with a standardized but protracted stimulus of 60 s (see Sec. I).

Differences between males in the sounds produced during the inter-grunt periods are reflected in differences in the BW_{20} (Fig. 2C). Some males produce relatively *pure* moans,

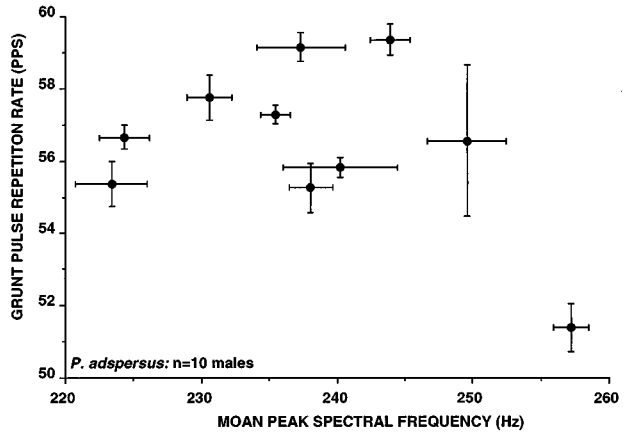


FIG. 5. Individual differences for 10 individual *P. adspersus*. Each individual is represented by its mean grunt pulse repetition rate (PPS) and the peak frequency of the moan. The standard error of the means are provided, and the sample size in each case is eight sounds. Individual males differ significantly in these as well as several other acoustic parameters (see text).

with sharply peaked spectra and small BW_{20}'s, for the duration of a courtship interaction. Other individuals produce moans with broader spectra and relatively large BW_{20}'s. During prolonged interactions, males may initially generate relatively pure moans but then appear to fatigue and grade into a signal that is closer to the growl and has a broader spectrum.

A few males were able to continually generate pure moans ($BW_{20} \leqslant 20$ Hz) for the duration of the 60-s stimulus (Fig. 8A, Fig. 9A), but for most the BW_{20}'s began to increase after 15–20 s. A few individuals appeared incapable of generating pure moans, even early in the stimulus presentation trial (Fig. 8B, Fig. 9B). The rate of moan production ($BW_{20} < 25$ Hz) varied from about 5 per min to 25 per min for the 7 individuals tested under this condition (Fig. 9C).

The purest moans (e.g., arrow A1) for the male illustrated in the sonagram of Fig. 8A appear as solid horizontal bands, had BW_{20}'s of about 10 Hz (Fig. 9A), and had nearly sinusoidal waveforms (Fig. 10A1). Later in the stimulus this individual began to grade in and out of moaning (arrows A2, A3 in Fig. 8A); the corresponding waveforms are illustrated in Fig. 10 (A2, A3). The strongly amplitude modulated signal produced in the middle of Fig. 10A3 is essentially a growl. Some individuals graded smoothly from a pure moan to a growl over the course of about a second (e.g., Fig. 10A4).

We suspect that the moan may be produced by continuous contraction of a sonic muscle at a rate equivalent to the fundamental frequency, and that it may be the most physically demanding sound to produce. Since this kind of sonic mechanism is expected to be highly temperature dependent, we examined the effects of temperature on sound production.

D. Temperature dependence of acoustic features

Moan peak frequency and grunt pulse repetition rate both increased systematically with temperature (Fig. 11), and the physiological Q_{10}'s governing these processes were indistinguishable (1.61 ± 0.06 SE for moan peak frequency and 1.75 ± 0.08 SE for grunt pulse repetition rate; 26–29 °C range; 5 males). Over the narrow temperature range evaluated, the relationship between these acoustic variables and temperature was quite linear, with correlation coefficients (R) of 0.67 ± 0.07 SE for moans and 0.74 ± 0.08 SE for grunts. The mean slope for moans was 10.1 Hz/° C ± 2 SE, and 3.3 pulses/s per °C ± 0.3 SE for grunts.

III. DISCUSSION

A. Species differences and the evolution of acoustic displays

The data presented here have allowed the first comparison of the acoustic displays of closely related mormyrid electric fish and have revealed clear species differences. The initial discovery of unusually elaborate male acoustic displays in a single species (*P. adspersus*) suggested that sounds might play a role in species and possibly even individual recognition (Crawford *et al.*, 1986). Many of the conspicuous social displays of animals are thought to have evolved through mechanisms of sexual selection (Darwin, 1871; Fisher, 1930), and inter-sexual-selection is widely believed to be a critical mechanism for the evolution of the displays presented by males to females during reproductive behavior; active female choice may provide a *selective force* that drives the elaboration of male displays (see Searcy and Andersson, 1986; Ryan, 1990; Ryan and Keddy-Hector, 1992). Since species recognition is fundamental to successful mating, it is expected that choice and differential reproductive success will drive the divergence of displays produced by closely related species. Our data, showing the acoustic separation of these two closely related species of electric fishes, provide additional support for the hypothesis that these signals function in mate choice. Acoustic species recognition in fishes, based on temporal patterning of acoustic pulses, has previously been demonstrated in the marine damselfishes (Pomacentridae; Myrberg *et al.*, 1978; Spanier, 1979), and freshwater sunfishes (Centrachidae; Gerald, 1971), and is likely a major function of sound production among mormyrids and the many other sound-generating fishes (reviewed by Fine *et al.*, 1977; Schellart and Popper, 1992; Myrberg, 1997).

B. Utility of individual differences

In theory, female *Pollimyrus* should benefit not only by choosing mates of the same species but also by choosing individuals of the highest ''quality.'' Females invest more heavily in the production of gemetes than males, and are consequently significantly larger than males by the time mating takes place. Additionally, field studies of *P. isidori* have shown that the sex ratio among breeding adults is skewed toward males, with about twice as many males as females (Crawford *et al.*, 1997). Under these circumstances, choice mechanisms for optimizing mate quality are expected to be particularly advantageous to females (see Emlen and Oring, 1977). Females could improve the yield on their reproductive investment by making informed choices from among the abundant males. In the laboratory, female *P. adspersus* often mate exclusively with a single male when give a choice be-

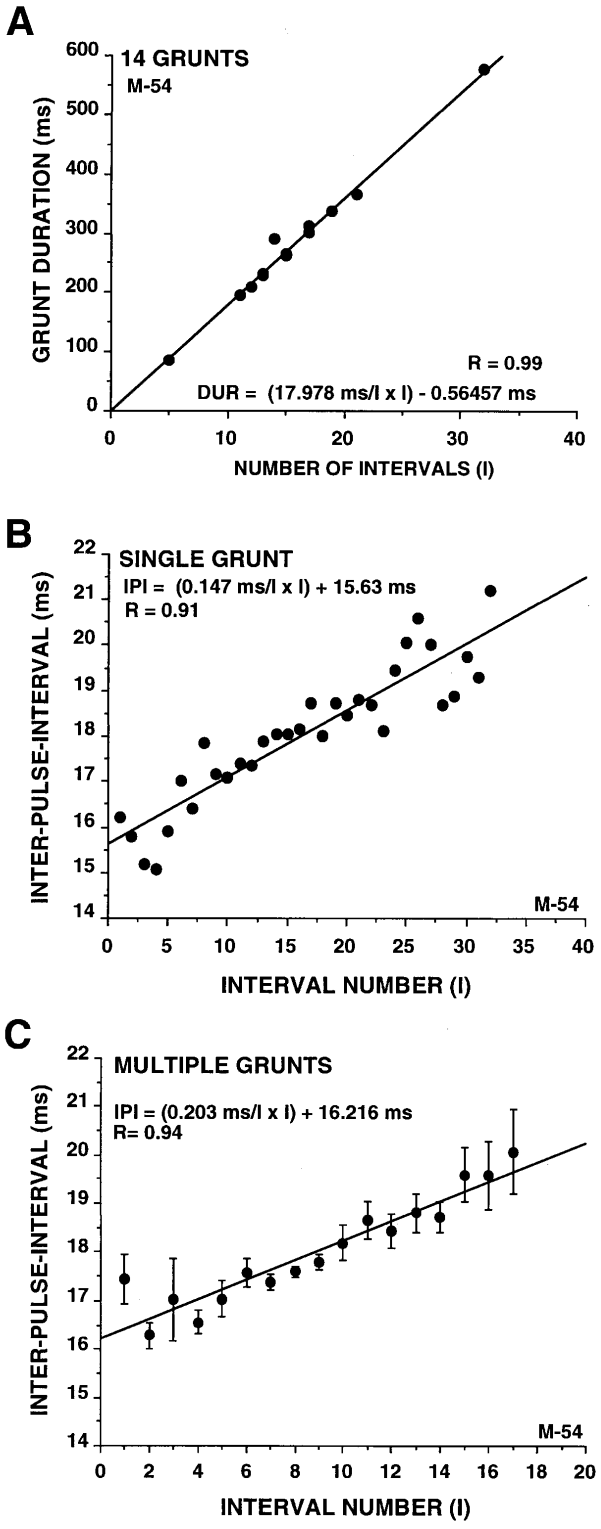


FIG. 6. Males tend to hold inter-pulse-interval relatively constant but can substantially change grunt duration by varying the number of intervals used. The top panel (A) shows the range of grunt durations as a function of interval number for a single individual. Most grunts for this individual fell in the 200–350 ms range, requiring about 11–19 intervals (12–20 pulses). Within a grunt (B), the duration of each successive interval increases slightly (≈150 μs/interval). Although small, this steady increase is a consistent feature of grunts made by any given individual (e.g., C). In C, the mean interval duration is plotted together with the SE. Fourteen grunts were used to calculate these means. Since the highest interval numbers only occurred in the longest grunts, the number intervals used for each mean declined with interval number (e.g., only 5 grunts had as many as 17 intervals).

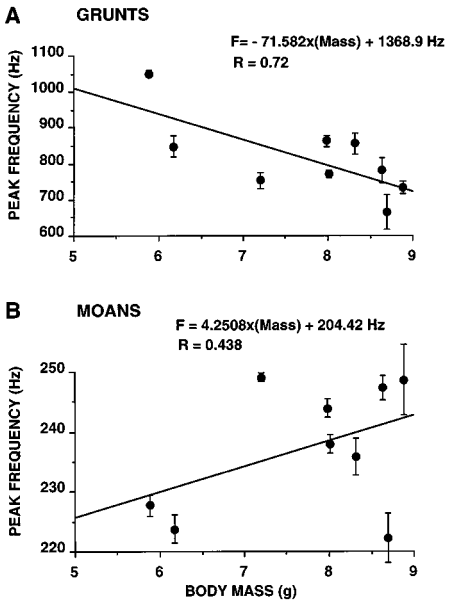


FIG. 7. Larger males produce grunts with a lower peak frequency (A). Peak frequency declines at about 72 Hz per gram body mass for these nine males. Moan peak frequency was statistically independent of body mass (B). Each individual is represented by its mean ±SE, eight sounds per mean.

tween two territory holders (pers obs). The basis of these choices is not yet understood but may involve acoustic cues.

Our analyses show that there are a variety of significant individual differences in acoustic characteristics among male *P. adspersus* which may provide a basis for decision making behavior. Field studies have demonstrated the importance of acoustically mediated individual recognition among other fish species. For example, male damselfishes recognize their conspecific neighbors by acoustic cues (Myrberg and Riggio, 1985), and female damselfishes use sounds in the selection of male mates (Myrberg *et al.*, 1986).

There are several ways in which acoustic information might be exploited by female *P. adspersus* for mate choice in their natural nocturnal environment. First, if some constellation of acoustic cues were to provide a ''signature'' for each of a group of available males, females could then associate each signature with other information about male quality derived through nonauditory mechanisms. Breeding males are territorial, building nests on their territories and providing exclusive care of the eggs and larvae after spawning. Females make repeated visits to available territories prior to spawning and it is during these visits that females experience the most vigorous sonic courtship by males. These visits occur in complete darkness, but females could use electrosensory and other sensory systems to discern characteristics of the territory. After visiting the available territories, a female could then use acoustic signals, detectable at a distance, to return to the best territory.

A second potential use of the individual differences in male acoustic behavior is for the assessment of male physical vigor. There are several ways in which vigor might impact female reproductive success. The success of breeding efforts depends in part on the ability of males to repel nest predators and conspecific rivals from their territories. In addition, both

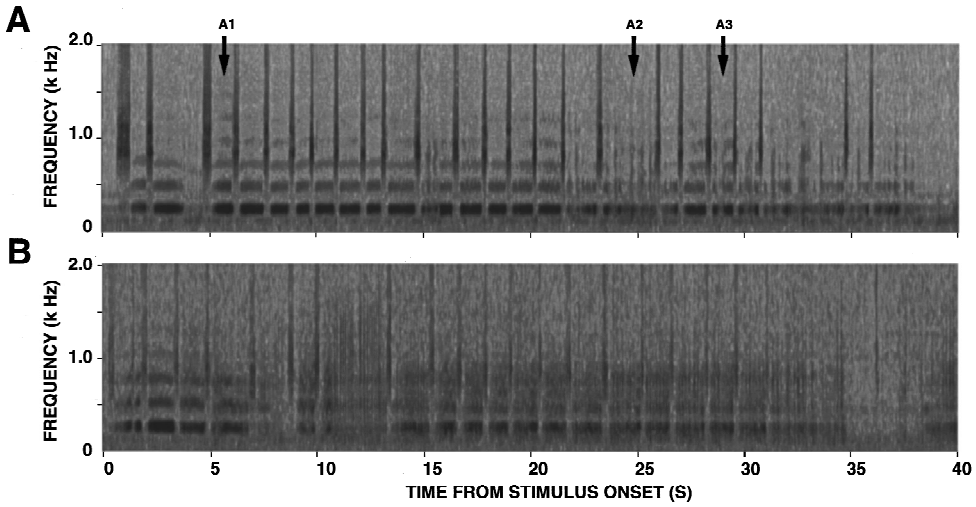


FIG. 8. Example responses during electric play-backs to two different males (*probe trials*). Each panel (A, B) shows a sonagram for the first 40 s of a 60-s stimulation period, illustrating the time course of the acoustic response. The full 60-s period is summarized in Fig. 9, for each of these two cases. Note that the male shown in A produced relatively narrow-band moans (horizontal bands between vertical grunts) out to about 22 s, and then appeared to fatigue. Examples of moans (arrow A1) and broader band signals (arrows A2, A3) are illustrated in Fig. 10. The second male (B) produced only a few narrow-band moans, and then graded into broadband signals that sound like growling. The BW_{20}'s corresponding to these two individuals are summarized in Fig. 9. Sonagrams are based on 1024-point FFTs, 512-point frames, and 75% overlap with a Hamming window.

males and females must be strong enough to endure a migration from the dry season river habitat to the flood plains for breeding. Insofar as vigorous males tend to father vigorous offspring, the representation of a female's genes in the future breeding populations may be enhanced by choosing vigorous mates. Since sound production is an energetically demanding biomechanical activity (see Crawford *et al.*, 1997), males that are able to produce and sustain ''robust'' displays may be in relatively good physical condition. The results of our experiments with simulated female visits show that there are substantial differences between males in their ability to produce the sonic courtship display, and this could be exploited by females during mate choice.

Our finding that there was a significant inverse correlation between grunt peak frequency and male mass raises the possibility that females exploit acoustic behavior for the direct measurement of mass. Larger males produce deeper grunts (i.e., lower peak frequency) than their diminutive competitors, and mass is likely to correlate with a number of the issues discussed above (e.g., vigor). Fortunately for females, the truthful advertisement of physical features may be an inescapable consequence of the mechanics of sound production. A dependence of spectral peak frequency on size has been shown in a variety of fish species (e.g., Stabentheiner, 1988; Laddich *et al.*, 1992; Myrberg *et al.*, 1993).

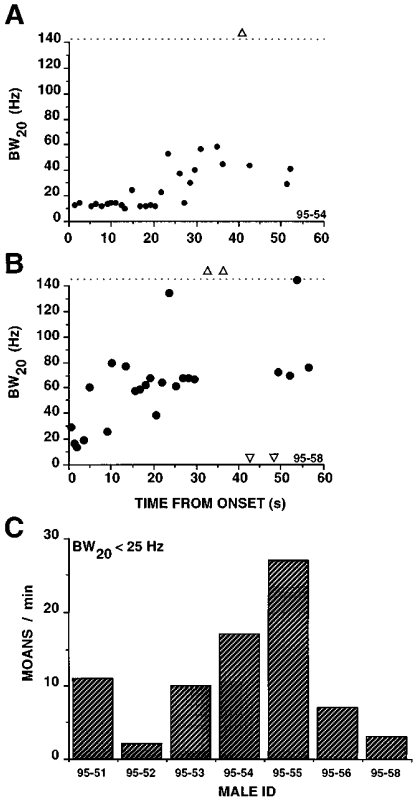


FIG. 9. Individual differences in moan *purity* (BW_{20}) during prolong electrical playbacks. Panels A and B show the same two individuals shown in the preceding figure (Fig. 8). Note the first individual (A) produced many signals with BW_{20} less than 25 Hz, whereas the second (B) produced only a few in the early part of the simulated female visit. Moan rates (BW_{20} <25 Hz) differed markedly among seven males tested under identical conditions (C). BW_{20} were measured from 16 384-point FFTs; 8196 point frame length; 87.5% overlap; Hamming window; 24-kHz sampling, 16 bits. Note upward triangles above the horizontal dashed lines (in A and B) represent BW_{20}'s>140 Hz, and downward triangles, along the horizontal axis, correspond to inter-grunt periods when there was no sound production (B only).

C. Auditory discrimination and communication

Additional behavioral experiments are needed to determine which of the acoustic cues discussed here are perceived and used by *Pollimyrus*. However, we do have some auditory neurophysiological data on *Pollimyrus* (Crawford, 1993, 1996) as well as behavioral data for another of the Mormyridae (*Gnathonemus petersii*; McCormick and Popper, 1984), and considerable behavior and neurophysiological

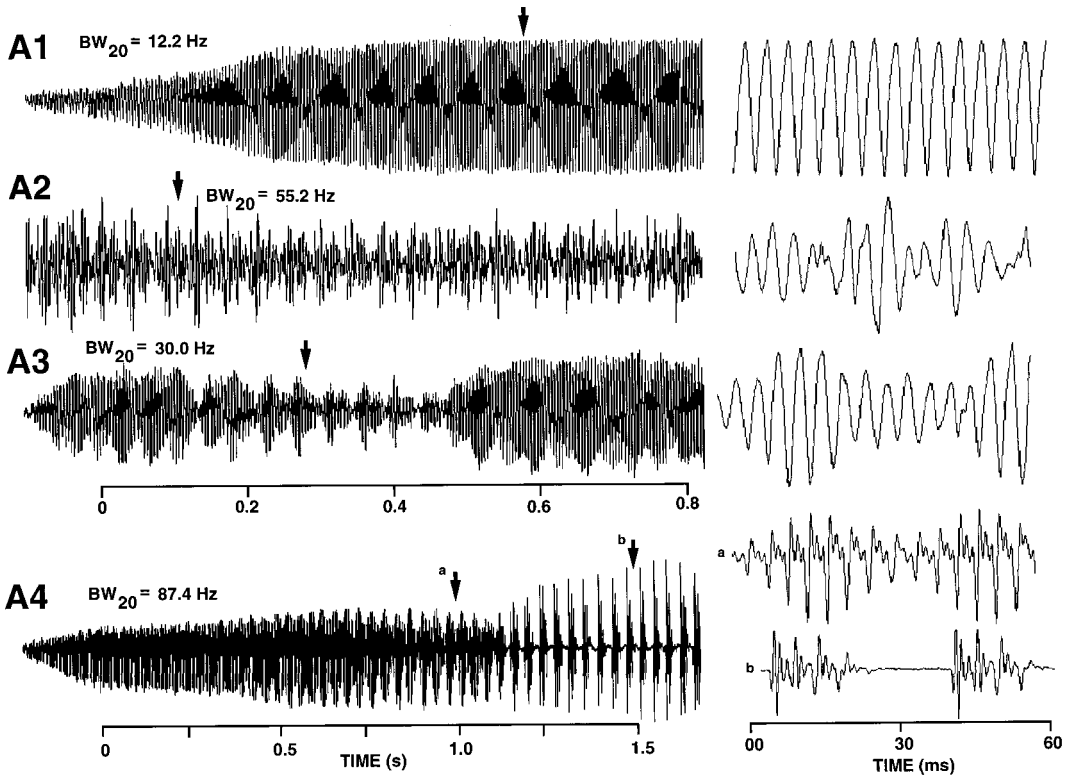


FIG. 10. Examples of moans (A1) and other inter-grunt signals produced by males during prolonged evoked courtship responses (A2–A4). In each case, the BW_{20} for the whole signal, and an expanded view of the waveform, are provided. The approximate time of the expanded waveform (right) is indicated by the vertical arrow(s) in each case. Panels A1–A3 correspond to the same fish shown in Fig. 8A, and A4 is from another individual.

data on hearing in the goldfish (*Carassius auratus*; see Furukawa and Ishii, 1967; Jacobs and Tavolga, 1968a; Popper, 1972; Fay, 1970, 1972, 1978a, 1978b, 1982, 1995; Fay and Passow, 1982; Coombs and Fay, 1985; Lu and Fay, 1993, 1995). The goldfish is also a teleost fish, but which has a different type of specialization for sound-pressure detection:

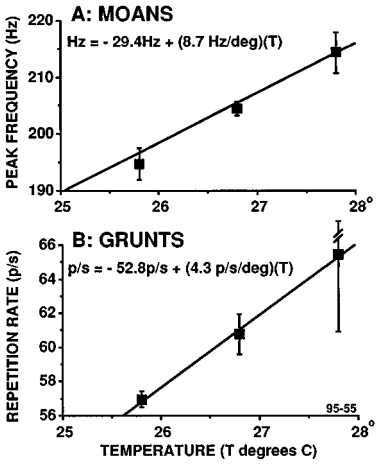


FIG. 11. Moan peak frequency (A) and grunt pulse repetition rate (B) both increased linearly with water temperature over the 26–29 °C range. The examples here are from an individual male *P. adspersus* examined at three different temperatures. The square points show the mean ±1 SE, based on seven or eight sounds at a given temperature. The lines are the least-squares best fit to the raw data, and the equation for each line are provided (T=temperature in °C; deg=°C; p/s=pulses per second; p/s/deg=pulses per second per °C).

A chain of ossicles (*Weberian ossicles*) couples a single large *swim bladder* in the abdominal cavity to the inner ear.

The neurophysiological data on *Pollimyrus* show that individual auditory neurons are highly sensitive to acoustic signals spanning the range of frequencies that dominates the communication signals they produce (Crawford, 1993). Behavioral studies of detection threshold in the mormyrid *G. petersii* also show high sensitivity to sounds over the same frequency range, 100–1500 Hz. Thus *Pollimyrus* and other mormyrids are clearly highly sensitive within the same frequency band used by the fish during sound production.

At the level of the auditory midbrain (mesencephalon) in *Pollimyrus*, many neurons exhibit selectivity for temporal periodicities characteristic of the sounds used in communication (Crawford, 1997). It has been hypothesized that these *tuned* responses arise through a neural coincidence detection mechanism, similar to the ''autocorrelationlike timing mechanism'' postulated by Fay and Passow (1982). In behavioral studies, Fay and Passow demonstrated that goldfish could discriminate small changes in the timing of acoustic pulses, with just discriminable differences (jnds) in interpulse-interval below 200 μs (at rates of 200 pulses per second); jnds in these and other discrimination tests scaled as a power function of stimulus period, with the best performance (smallest jnds) at the highest rates (i.e., shortest periods). They suggested that central neurons with the kind of selectivity demonstrated in *Pollimyrus* (Crawford, 1997) might underlie the sort of time-based discrimination behavior they discovered in goldfish. Together, these neurophysiological

results for mormyrids, and the behavioral findings of Fay and Passow, suggest that fish like *Pollimyrus* spp probably have high behavioral sensitivity to the small temporal differences in the sounds made by different species, and perhaps even the differences between individuals.

P. isidori and *P. adspersus* differed by 12 pps in the mean pulse repetition rate (pps) of grunts, with the midpoint between the two distributions at about 50 pps (Fig. 4A). To the degree that we can predict mormyrid behavior from experiments with goldfish (Fay and Passow, 1982), we would expect that grunts differing by 12 pps would be readily discriminated. When listening to pulse trains at 50 pps, goldfish detected jitter in the intervals comprising the pulse train when the standard deviation of the interval distribution was about 2 ms or more. Since a 2-ms change in interval (at 50 pps) corresponds to a change in pulse repetition rate of about 4.5 pps, we expect that the mean difference between the grunts of *P. isidori* and *P. adspersus* (12 pps) corresponds to roughly 3 jnd's, a substantial *perceptual* difference. Similarly, the range of pulse repetition rates used by different individual male *P. adspersus* is sufficiently large that at least some individual discrimination on the basis of this dimension is also likely (Fig. 5). It should be noted that the task of discriminating between two different grunts is not identical to any of the various psychophysical procedures used with goldfish, and that the jnd for the grunt discrimination task may be smaller than the estimated 4.5 pps. Field studies of a complex of four species of marine damselfishes (*Eupomacentrus*) have shown that males discriminate between con- and hetero-specific gruntlike sounds (*chirps*) that differ in pulse repetition rate and pulse number; pulse repetition rate is one of several cues that is used for species recognition and the four species span a range of only about 10 pps (Spanier, 1979).

The moans produced by *P. isidori* and *P. adspersus* are also likely to be perceptually distinct, based on studies of discrimination in goldfish (Fay, 1970, 1978a, 1978b). Moans (Fig. 2C, D) are relatively simple sounds that usually have two main, harmonically related, components. The maximum peak of the amplitude spectrum corresponds to the period of the signal's waveform. The separation between the means of the moan frequency distributions is 92 Hz, with the midpoint between the two species at about 286 Hz (Fig. 4B). In pure-tone frequency discrimination tasks, the jnd's for goldfish are usually about 3%–5% of the *neutral* stimulus frequency (Fay, 1970). Assuming similar discrimination in mormyrids, we would then expect jnd's of about 10 Hz and that on average the moans of our two species would be separated perceptually by about 9 jnd's. The range of individual differences within a selected group of 10 *P. adspersus* spans about 50 Hz (Fig. 5), and again at least some pairs of individuals would probably be readily distinguished based on moan period. Additionally, there was a large species difference in moan duration (Fig. 4D: 121 vs 821 ms), and this difference is also likely to be detectable by these fish. To our knowledge duration sensitivity has not yet been directly measured in behavioral studies of fishes, but small changes (a few percent) in signal duration are readily detected by other vertebrate animals (reviewed in Fay, 1988), and duration is known to have a pronounced influence on signal detectability (Popper, 1972; Fay and Coombs, 1983) and intensity discrimination (Fay and Coombs, 1992) in other fish (*Carassius auratus*).

Clearly psychophysical data on *Pollimyrus* are needed since there may well be important differences between mormyrids and goldfish. Additionally, though the psychophysical studies of goldfish are elegant and well conceived they were not designed to specifically address questions about discrimination of features of communication sounds. Further refinements in the estimates of jnds may follow from experiments modeled after natural behavioral tasks. It is also worth noting that the above discussion takes one stimulus dimension at a time, asking questions about perception and discrimination. Under normal circumstances the animals will have access to multiple acoustic cues that can potentially be used in parallel for discrimination (see Doherty, 1985), and recent generalization studies indicate that complex sounds present multiple perceptual cues to goldfish (Fay, 1995).

D. Mechanisms of sound production

The effect of temperature on the features of *P. adspersus* sounds is consistent with a sonic mechanism in which sound structure is influenced by the rate and force of contraction of a sonic muscle. Muscle-based drumming mechanisms are known in a variety of fish species (e.g., Cohen and Winn, 1967; Demski *et al.*, 1973), and the specific influence of temperature on rate in *P. adspersus* (grunt PRR and moan F_1) was quantitatively similar to that observed in other systems where the muscle-based sound production is better understood. For example, the fundamental of the toadfish (*Opsanus tau*) *boatwhistle* increases with slope (7.5 Hz/deg) and Q_{10} (1.5 in 24–26 °C range) that are close to what we have reported here for *P. adspersus* (based on Fine, 1978, Fig. 7; see also Brantly and Bass, 1994). Recent studies have revealed that male *P. adspersus* have a large muscle enveloping the posterior pole of the swimbladder, and local anesthesia of this muscle renders males temporarily mute (Huang *et al.*, 1996). We suspect that the moan is produced by repeated contraction of this muscle, and that the peak frequency of the moan spectrum corresponds to the rate of contraction (see Skoglund, 1961; Fine, 1978). Our finding that moan spectral peak is independent of male mass is consistent with this idea.

The *P. adspersus* grunt appears to be produced by a series of relatively widely spaced contractions producing the sequence of impulses sounds (pulses) which constitute this sound. The spectrum of the whole grunt in this species was almost entirely determined by the spectrum of the constituent pulses. We suggest that the grunt spectrum may thus provide a rough indication of the impulse response of the fish, and that this is largely determined by the resonance properties of the swimbladder (see discussion by Myrberg *et al.*, 1993). This hypothesis is supported by the observed systematic decline in the spectral peak of the grunt with male size; bladder volume scales with body volume in this species (unpub obs).

Both modeling (e.g., Love, 1978) and empirical studies (e.g., Løvik and Hovem, 1979) indicate that bladder resonances for small fish like *P. adspersus* in very shallow water

($\approx$0.5 m) should be in the approximate range we observed (i.e., 650–1050 Hz). Nevertheless, resonance depends on several variables that are, at present, unknown for *P. adspersus*, for example, the surface tension of the bladder. Similar relationships between peak frequency and size have been demonstrated in damselfish (Myrberg *et al.*, 1993; Lobel and Mann, 1995). Over the narrow range of masses we analyzed, our data fit equally well a power function [$F_{\text{peak}} = (2818\ \text{Hz/g}^{-0.617})*(\text{Mass})^{-0.617}$] and a linear function (slope$=-72$ Hz/g). This linear rate of change is very close to that reported for damselfish of comparable size by Myrberg *et al.* (approximately -60 Hz/g in the 4–10 g range, Fig. 3 of Myrberg *et al.*, 1993).

The mormyrid fishes are nocturnal and have developed sensory systems allowing them to penetrate an essentially aphotic ecological niche. Compared with other fishes, the acoustic repertoires of these mormyrids are elaborate. The sounds contain a great deal of information of potential significance for communication. Future research is being directed at uncovering which parameters are exploited by the fish in natural behavior, and how the information is analyzed in the auditory nervous system.

ACKNOWLEDGMENTS

We are most appreciative of the assistance with acoustic analysis provided by Katherine Engstrom, Camille A. Henry, and Xiaofeng Huang. James Kozloski, Virginia Richards, Robert Seyfarth, and Dirk Trojan provided valuable critical input during the research. We are indebted to Philippe Jacob and Vincent Bénech for their collaboration during work in Mali, and to the three anonymous referees who provided thoughtful critical comments on the manuscript. This research was supported by a fellowship from the University of Pennsylvania Scholars Foundation (ASH), and by grants from the National Institute of Deafness and Other Communication Disorder (NIH DC01252) and the PA Lions Hearing Research Foundation.

Bigorne, R. (**1990**). ''Revision systematique du genre *Pollimyrus* (Teleostei, Mormyridae) en Afrique de l'Ouest,'' Rev. Hydrobiol. **4**, 313–327.

Brantley, R. K., and Bass, A. H. (**1994**). ''Alternative male spawning tactics and acoustic signals in the plainfin midshipman *Porichthys notatus* (Teleostei, Batrachoididae), '' Ethology **96**, 213–232.

Bratton, B. O., and Kramer, B. (**1989**). ''Patterns of the electric discharge during courtship and spawning in the mormyrid fish, *Pollimyrus isidori*,'' Behav. Ecol. Sociobiol. **24**, 349–368.

Cohen, M. J., and Winn, H. E. (**1967**). ''Electrophysiological observations on hearing and sound production in the fish *Porichthys notatus*,'' J. Exp. Zool. **165,** 355–370.

Coombs, S., and Fay, R. R. (**1985**). ''Adaptation effects on amplitude modulation detection: behavioral and neurophysiological assessment in the goldfish auditory system,'' Hearing Res. **19,** 57–71.

Coombs, S., and Popper, A. N. (**1982**). ''Structure and function of the auditory system in the clown knifefish, *Notopterus chitala*, '' J. Exp. Biol. **97,** 225–239.

Crawford, J. D. (**1991**). ''Sex recognition by electrical cues in a sound producing mormyrid fish, *Pollimyrus isidori*,'' Brain Behav. Evol. **38**, 20–38.

Crawford, J. D. (**1992**). ''Individual and sex specificity in the electric organ discharges of breeding mormyrid fish, *Pollimyrus isidori*,'' J. Exp. Biol. **164,** 79–102.

Crawford, J. D. (**1993**). ''Central auditory neurophysiology of a sound-producing mormyrid fish: The mesencephalon of *Pollimyrus isidori*,'' J. Comp. Physiol. **172,** 1–14.

Crawford, J. D. (**1995**). ''Communication sounds of the West African electric fish *Pollimyrus isidori* recorded in the field: a description and comparison with *P. adspersus*,'' in Midwinter Meeting, Assoc. Res. Otolaryngol., St. Petersburg, FL **13**, 260.

Crawford, J. D. (**1996**). ''Feature detection by auditory neurons in the brain of a sound-producing fish,'' J. Comp. Physiol. A **180**, 439–450.

Crawford, J. D. (**1997**). ''Hearing and acoustic communication in the mormyrid electric fishes,'' Mar. Freshwater Behav. Physiol. **29** (in press).

Crawford, J. D., Cook, A. P., and Herberlein, A. H. (**1995**). ''Species differences in the acoustic signals of congeneric electric fishes: *Pollimyrus isidori and P. adspersus*,'' in *Neural Mechanisms of Behavior, Proceedings of the 4th International Congress on Neuroethology*, Cambridge, U.K. (Georg Thiem Verlag, Stuttgart, Germany), p. 350.

Crawford, J. D., Hagedorn, M. M., and Hopkins, C. D. (**1986**). ''Acoustic communication in an electric fish, *Pollimyrus isidori* (Mormyridae),'' J. Comp. Physiol. A **159,** 297–310.

Crawford, J. D., Jacobe, P., and Benech, V. (**1977**). ''Field studies of a strongly acoustic fish in West Africa: reproductive ecology and acoustic behavior of *Pollimyrus isidori*, MORMYRIDAE,'' Behaviour **134**, 1–49.

Darwin, C. (**1871**). *The Descent of Man and Selection in Relation to Sex* (Appleton, New York).

Demski, L. S., Gerald, J. W., and Popper, A. N. (**1973**). ''Central and peripheral mechanisms of teleost sound production,'' Am. Zool. **13**, 1141–1167.

Doherty, J. (**1985**). ''Trade-off phenomena in calling song recognition and phonotaxis in the cricket, *Gryllus bimaculatus* (Orthoptera, Gryllidae),'' J. Comp. Physiol. A **156,** 787–801.

Emlen, S. T., and Oring, L. W. (**1977**). ''Ecology, sexual selection, and the evolution of mating systems,'' Science **197,** 215–223.

Fay, R. R. (**1970**). ''Auditory frequency discrimination in the goldfish (*Carassius auratus)*,'' J. Comp. Physiol. Psychol. **73,** 175–180.

Fay, R. R. (**1972**). ''Perception of amplitude-modulated auditory signals by the goldfish,'' J. Acoust. Soc. Am. **52,** 660–666.

Fay, R. R. (**1978a**). ''Phase-locking in goldfish saccular nerve fibers accounts for frequency discrimination capacities,'' Nature (London) **275,** 320–322.

Fay, R. R. (**1978b**). ''Coding of information in single auditory-nerve fibers of the goldfish,'' J. Acoust. Soc. Am. **63,** 136–146.

Fay, R. R. (**1982**). ''Neural Mechanisms of an Auditory Temporal Discrimination by the Goldfish,'' J. Comp. Physiol. A **147,** 201–216.

Fay, R. R. (**1988**). *Hearing in Vertebrates: A Psychophysics Databook* (Hill-Fay Associates, Winnetka, Illinois).

Fay, R. R. (**1995**). ''Perception of spectrally and temporally complex sounds by the goldfish (*Carassius auratus*),'' Hearing Res. **89,** 146–154.

Fay, R. R., and Coombs, S. (**1983**). ''Neural mechanisms in sound detection and temporal summation,'' Hearing Res. **10,** 69–92.

Fay, R. R., and Combs, S. (**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.

Fay, R. R., and Passow, B. (**1982**). ''Temporal discrimination in the goldfish,'' J. Acoust. Soc. Am. **72,** 753–760.

Fay, R. R., Hillery, C. M., and Bolan, K. (**1982**). ''Representation of sound pressure and particle motion information in the midbrain of the goldfish,'' Comp. Biochem. Physiol. **71**, 181–191.

Fine, M. L. (**1978**). ''Seasonal and geographical variation of the mating call of the oyster toadfish, *Opsanus tau L.*,'' Oecologia **36**, 45–57.

Fine, M. L., Winn, H. E., and Olla, B. L. (**1977**). ''Communication in fishes,'' in *How Animals Communicate* (Indiana U. P., Bloomington), pp. 472–518.

Fisher, R. A. (**1930**). *The Genetical Theory of Natural Selection* (Clarendon, Oxford, U.K.).

Furukawa, T., and Ishii, Y. (**1967**). ''Neuorophysiological studies on hearing in goldfish,'' J. Neurophysiol. **30,** 1377–13403.

Gerald, J. W. (**1971**). ''Sound production during courtship in six species of sunfish (Centrarchidae),'' Evolution **25**, 75–87.

Huang, X., Kozloski, J., and Crawford, J. D. (**1996**). ''Sexually dimorphic swimbladder muscles in the sonic fish *Pollimyrus isidori* (Mormyridae),'' Soc. Neurosci. **22**, 446.

Jacobs, D., and Tavolga, W. (**1968**). ''Acoustic frequency discrimination in the goldfish,'' Animal Behav. **16,** 67–71.

Kirschbaum, F. (**1987**). ''Reproduction and development of the weakly elec-

tric fish, *Pollimyrus isidori* (Mormyridae, Teleostei) in captivity,'' Environ. Biol. Fishes **20**, 11–31.
Laddich, F., Brittinger, W., and Kratochvil, H. (**1992**). ''Significance of agonistic vocalization in the croaking gourami (*Trichopsis vittatus*, Teleostei),'' Ethology **90**, 307–314.
Lobel, P. S., and Mann, D. A. (**1995**). ''Spawning sounds of the damselfish, *Dascyllus albisella* (Pomacentridae), and relationship to male size,'' Bioacoustics **6**, 187–198.
Love, R. (**1978**). ''Resonant acoustic scattering by swimbladder-bearing fish,'' J. Acoust. Soc. Am. **64,** 571–580.
Lovik, A., and Hovem, J. M. (**1979**). ''An experimental investigation of swimbladder resonance in fishes,'' J. Acoust. Soc. Am. **66,** 850–854.
Lu, Z., and Fay, R. (**1993**). ''Acoustic response properties of single units in the torus semicircularis of the goldfish, *Carassius auratus*,'' J. Comp. Physiol. A **173,** 33–48.
Lu, Z., and Fay, R. (**1995**). ''Acoustic response properties of single neurons in the central posterior nucleus of the thalamus of the goldfish, *Carassius auratus*,'' J. Comp. Physiol. A **176,** 747–760.
McCormick, C. A. (**1992**). ''Evolution of central auditory pathways in anamniotes,'' in *The Evolutionary Biology of Hearing* (Springer-Verlag, New York), pp. 323–350.
McCormick, C. A., and Popper, A. N. (**1984**). ''Auditory sensitivity and psychophysical tuning curves in the elephant nose fish,'' J. Comp. Physiol. **155,** 753–761.
Myrberg, A. A., Jr. (**1997**). ''Underwater sound: expanding relevance to behavioral function,'' Mar. Freshwater Behav. Physiol. **29** (in press).
Myrberg, A. A., Jr., and Riggio, R. J. (**1985**). ''Acoustically mediated individual recognition by a coral reef fish (*Pomacentrus partitus)*,'' Animal Behav. **33,** 411–416.
Myrberg, Jr., A. A., Ha, S. J., and Shamblott, M. J. (**1993**). ''The sounds of bicolor damselfish (*Pomacentrus partitus*): Predictors of body size and a spectral basis for individual recognition and assessment,'' J. Acoust. Soc. Am. **94,** 3067–3070.
Myrberg, Jr., A. A., Mohler, M., and Catala, J. C. (**1986**). ''Sound production by males of a coral reef fish (*Pomacentrus partitus*): Its significance to females,'' Animal Behav. **34,** 923–933.
Myrberg, A. A., Jr., Spanier, E., and Ha, S. J. (**1978**). ''Temporal patterning in acoustical communication,'' in *Contrasts in Behavior* (Wiley, New York), pp. 136–179.
Popper, A. (**1972**). ''Auditory thresholds of the goldfish as a function of signal duration,'' J. Acoust. Soc. Am. **52,** 596–602.
Popper, A. N., and Fay, R. R. (**1984**). ''Sound detection and processing by teleost fish: a selective review,'' in *Comparative Physiology of Sensory Systems* (Cambridge U. P., Cambridge), pp. 67–101.
Popper, A. N., Saidel, W. M., and Chang, J. S. Y. (**1993**). ''Two types of sensory hair cell in the saccule of a teleost fish,'' Hearing Res. **64,** 211–216.
Ryan, M. J. (**1990**). ''Signals, species and sexual selection,'' Am. Sci. **8,** 46–52.
Ryan, M. J., and Keddy-Hector, A. (**1992**). ''Directional patterns of female mate choice and the role of sensory bias,'' Am. Nat. **139,** S4–S35.
Saidel, W., Lanford, P., Yan, H., and Popper, A. (**1995**). ''Hair cell heterogeneity in the goldfish saccule,'' Brain Behav. Evol. **46**, 362–370.
Schellart, N. A. M., and Popper, A. N. (**1992**). ''Functional aspects of the evolution of the auditory system of actinopterygian fish,'' in *The Evolutionary Biology of Hearing* (Springer-Verlag, New York), pp. 295–322.
Schmidt-Nielsen, K. (**1979**). *Animal Physiology: Adaptation and Environment* (Cambridge U. P., Cambridge, U.K.), pp. 207–209.
Searcy, W. A., and Andersson, M. (**1986**). ''Sexual selection and the evolution of song,''Ann. Rev. Ecol. Syst. **17,** 507–533.
Skoglund, C. R. (**1961**). ''Functional analysis of swimbladder muscles engaged in sound production of the toadfish,'' J. Biophys. Cytol. Suppl. **10**, 187–200.
Spanier, E. (**1979**). ''Aspect of species recognition by sound in four species of damselfishes, genus Eupomacentrus (Pisces: Pomacentridae),'' Z. Tierspsychol. **51**, 301–316.
Stabentheiner, A. (**1988**). ''Correlations between hearing and sound production in piranhas,'' J. Comp. Physiol. A **162,** 67–76.
Stipetic, E. (**1939**). ''Über das Gehörorgan der Mormyriden,'' Z. Vergl. Physiol. **26**, 740–752.

Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques

W. Tecumseh Fitch
Program in Speech and Hearing Sciences and Department of Anthropology, 33 Kirkland Street, Room 1036, Harvard University, Cambridge, Massachusetts 02138

(Received 11 December 1996; accepted for publication 14 April 1997)

Body weight, length, and vocal tract length were measured for 23 rhesus macaques (*Macaca mulatta*) of various sizes using radiographs and computer graphic techniques. Linear predictive coding analysis of tape-recorded threat vocalizations was used to determine vocal tract resonance frequencies (''formants'') for the same animals. A new acoustic variable is proposed, ''formant dispersion,'' which should theoretically depend upon vocal tract length. Formant dispersion is the averaged difference between successive formant frequencies, and was found to be closely tied to both vocal tract length and body size. Despite the common claim that voice fundamental frequency (F_0)provides an acoustic indication of body size, repeated investigations have failed to support such a relationship in many vertebrate species including humans. Formant dispersion, unlike voice pitch, is proposed to be a reliable predictor of body size in macaques, and probably many other species. [S0001-4966(97)01808-0]

PACS numbers: 43.80.Ka, 43.80.Jz, 43.70.Aj, 43.64.Bt [FD]

INTRODUCTION

The acoustic characteristics of animal vocalizations have been postulated to provide information about many important attributes of the vocalizer, including its size, age, sex, and reproductive status and emotional state. A particularly important set of parameters may relate to the size of the vocalizer. Body size is an extremely important variable in animal physiology, mortality, and ecology (Peters, 1983; Schmidt-Nielsen, 1984; Harvey, 1990; Alexander, 1996), and in many types of animal social behavior, including most prominently aggressive interactions (Parker, 1974; Clutton-Brock and Albon, 1979) and mating behavior (Clutton-Brock *et al.*, 1977). The size of various components of the sound production apparatus (e.g., the lungs, vocal folds, and vocal tract) has an important effect on the acoustic output (Fant, 1960; Lieberman, 1984). Because the size of these production components may in many cases be related to the overall weight or length of the animal, there is good reason to expect that some aspects of the acoustic signal may provide cues to the size of the vocalizer.

The most frequently cited acoustic parameter which could provide a cue to body size is mean and/or lowest fundamental frequency (Darwin, 1871). In particular, the lowest producible fundamental frequency of phonation ($F_{0\min}$) is determined by the length and mass of the vocal folds: the larger the folds, the lower is $F_{0\min}$. If the size of the vocal folds were related to the vocalizer's body size, $F_{0\min}$would provide a good cue to body size (Morton, 1977; Hauser, 1993). This indeed appears to be the case in some species, including some toads and frogs (Martin, 1972; Davies and Halliday, 1978; Ryan, 1988). However, such a relationship between body size and vocal fold size does not seem to be typical in other vertebrates. For instance, there is no correlation between F_0 and body size in adult humans (Lass and Brown, 1978; Künzel, 1989; Cohen *et al.*, 1980; van Dommellen, 1993), red deer (McComb, 1991), and other amphibian species (Asquith and Altig, 1990; Sullivan, 1984). This lack of correlation in adult humans is particularly surprising given the widespread assumption that a ''deep'' or low-pitched voice indicates large body size.

The lack of correlation between F_0 and size seems less surprising when the anatomy of the vocal folds is considered. The folds are housed within the flexible cartilaginous larynx, which itself floats at the top of a trachea and is unconstrained in size by neighboring bony structures [the hyoid bone, though ossified, grows as a unit with the larynx, Schön (1971) and Schneider *et al.* (1967)]. Thus the larynx and vocal folds can grow independently of the rest of the head or body, as indeed occurs in human males at puberty (Negus, 1949; Goldstein, 1980). At puberty, androgen receptors in the laryngeal cartilages respond to increased circulating testosterone with a profound growth spurt (Tuohimaa *et al.*, 1981; Beckford *et al.*, 1985). The result is a typical F_0 for adult males which is about half that of adult females, despite an average difference in body weights of only 20% (Hollien, 1960). Hypertrophy of the male larynx, out of all proportion to body size, is carried to an absurd extreme in animals such as the howler monkey [*Allouatta seniculus*, Schön (1971)], in which the larynx and hyoid together are the size of the entire skull, or the hammerhead bat [*Hypsignathus monstrosus*, Kingdon (1974) and Schneider *et al.* (1967)], where the larynx of the male occupies virtually the entire thoracic cavity.

When such developmental flexibility is present there is clearly no *a priori* reason to expect vocal fold size (and thus F_0) to be well-correlated with body size (Fitch, 1994; Fitch and Hauser, 1995). Despite the common claim that voice pitch provides an accurate cue to body size (e.g., Morton, 1977), these data suggest that the larynx is ill-suited to provide dependable cues to body size.

A different potential acoustic cue to body size derives from the fact that, in most vertebrates, the sound signal cre-

ated in the larynx passes through the supralaryngeal vocal tract (hereafter, simply ''vocal tract'') before being radiated into the environment. The column of air in the vocal tract has certain natural modes of vibration or formants, which affect the resultant output signal (Fant, 1960) (the terminology of speech scientists is adopted here, using ''formants'' as opposed to the more general term ''resonances,'' despite the fact that the sounds produced by nonhumans differ in significant ways from human speech). If the cross-sectional area function of the vocal tract is constant, the primary determinant of formant frequencies is the length of the vocal tract (Fant, 1960; Lieberman and Blumstein, 1988). In particular, a lengthening of the vocal tract tube will lead to a decrease in the average spacing between successive formants, or ''formant dispersion.''

Several researchers in speech science have postulated a relationship between body size and formant frequencies (Fant, 1960; Nearey, 1978; Lieberman, 1984; Peterson and Barney, 1952), and Peterson and Barney provided evidence suggestive of such a relationship in humans (children have higher formant dispersion than adults, and women greater than men). Because the mammalian vocal tract is made up of the pharyngeal, oral, and nasal cavities, which are firmly bounded by the bones of the skull, and skull size is closely tied to overall body size (Morita and Ohtsuki, 1973; Dechow, 1983; Janis, 1990; Valkenburgh, 1990; Alcantara *et al.*, 1991; Sharma, 1990; and see data below), vocal tract length should be much less free to vary independently of body size than larynx size. Thus we can expect vocal tract length and the attendant acoustic cue of formant dispersion to provide a correspondingly more robust cue to body size.

Systematic investigation of the role of the supralaryngeal vocal tract in nonhuman vocalization was inaugurated with the work of Lieberman (Lieberman, 1968; Lieberman *et al.*, 1969; Lieberman, 1975; Lieberman, 1984), who was interested primarily in the differences between the human and nonhuman vocal tract. Lieberman showed that production of the full range of vowels, in particular the /i/ and /u/ vowels which are ubiquitous in human language, is impossible without the ''two-tube'' vocal tract of modern human beings. However, Lieberman (1968, p. 1576) also described modifications of monkey calls via changes in vocal tract length, and did not rule out the possibility that such changes have communicative significance. Further suggestions of a possible communicative role for vocal tract resonances in nonhumans came from spectrographic analyses of baboons (*Papio hamadryas*), geladas (*Theropithecus gelada*), and vervets (*Cercopithecus aethiops*) (Andrew, 1976; Richman, 1976; Seyfarth and Cheney, 1984, respectively).

More recent work has documented the role of supralaryngeal articulation in modifying acoustic characteristics of nonhuman vocalizations. Bauer (1987) built on Marler and Tenaza's (1977) technique of frame-by-frame video analysis of vocal production in chimpanzees (*Pan troglodytes*) to show that F_0 is related to mouth-opening size during chimp vocalizations. However, whether any causal acoustic mechanism underlies this correlation remains unclear. Hauser *et al.* (1993) used video analysis to analyze rhesus macaque calls, showing that changes in the first resonance frequency were well-correlated with changes in mandible position. A similar result was documented in cat (*Felis domesticus*) vocalizations by Shipley *et al.* (1991). Finally, Hauser and Schön-Ybarra (1994) experimentally eliminated vocal tract elongation via lip movements (using injections of the nerve-blocker xylocaine into the peri-oral region of rhesus macaques). They found that resonance frequencies were significantly higher than normal in the ''coo'' vocalization, which is normally accompanied by rounded lips. No changes were observed in F_0 or call duration. These experimental data are thus consistent with the predictions of source–filter theory (Fant, 1960), and indicate independence of source and filter in these calls.

Overall, a wealth of data suggests that the principles of source–filter theory and acoustic phonetics, originally developed for human speech, are applicable to nonhuman vocalizations as well. Source and filter appear to be independent in most cases, and movements of the articulators affect the vocal tract filter in the predicted ways. A variety of vertebrates can use the differences in formant frequency to discriminate synthesized vowels (baboons: Heinz and Brady, 1988; dogs: Baru, 1975; cats: Dewson, 1964; blackbirds and pigeons: Heinz *et al.*, 1981), and macaques can perceive formants with accuracy rivaling that of humans (Sommers *et al.*, 1992). However, the only conclusive evidence that supralaryngeal acoustic cues are utilized in nonhuman communication comes from Owren's work with vervet monkeys (*Cercopithecus aethiops*). Owren and Bernacki (1988) used linear predictive coding (LPC) analysis of vervet ''snake'' and ''eagle'' alarm calls to isolate characteristics of source waveform, presumed vocal tract filtering functions, and temporal patterning which distinguished these calls. Owren (1990b) then used an operant paradigm to test classification of synthetic calls in which each of these characteristics was modified independently. The results indicated that spectral characteristics played the dominant perceptual role in distinguishing the two call types, suggesting that supralaryngeal articulation conveys distinctive information in these calls.

The literature reviewed above suggests that vocal tract length and formant frequencies may be correlated with body size in many vertebrate species, and that nonhumans possess the perceptual machinery to make use of this correlation. However, despite the fact that several researchers have suggested such a relationship (Fant, 1960; Lieberman, 1984; Peterson and Barney, 1952), no study has addressed this issue directly in any species. The goal of this study was to measure body weight and body length, the lengths of the oral and nasal vocal tracts, and formant frequency dispersion in calls, and to quantify the relationship(s) between these variables, using rhesus macaques as the study species.

I. METHODS

A. Study animals

The subjects were 20 healthy rhesus macaques (*Macaca mulatta*), housed at the New England Regional Primate Research Center in Southboro, MA. Animals were between 1 and 9 years of age (mean 4.7 yr), and between 2.6 and 15.6

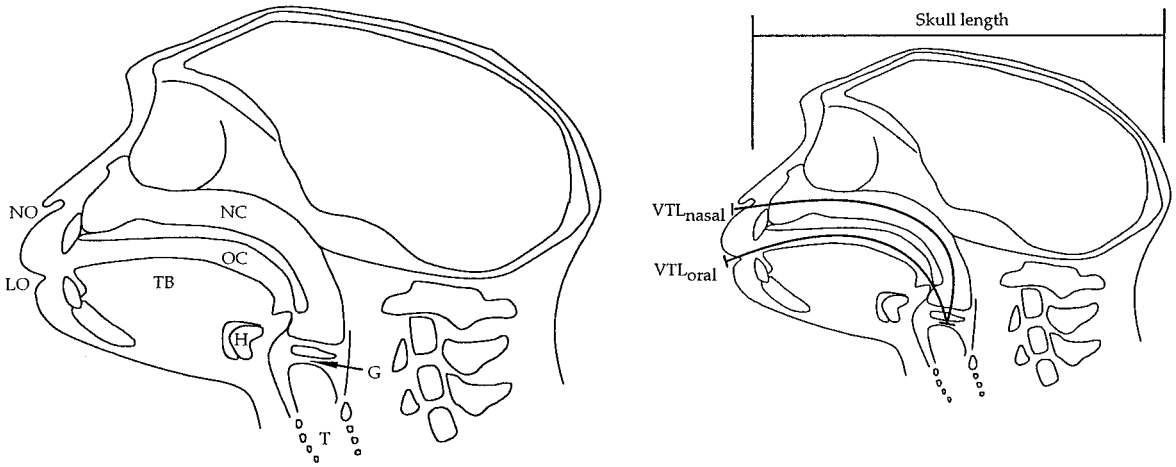


FIG. 1. Schematic of anatomical features (left) and morphometric features (right) used in this study: G: glottis, H: hyoid bone with subhyoid air sac, OC: oral cavity, LO: lip opening, NC: nasal cavity, NO: nostril opening, T: tracheal lumen, TB: tongue body, VTL: vocal tract length.

kg in weight (mean 7.13 kg). Both males ($N=11$) and females ($N=9$) were studied. Animals were fed primate chow, fresh fruit, and *ad libitum* water, and were maintained at their *ad libitum* body weight.

B. Anatomical measures

Animals were anesthetized with 5–7 mg/kg of Telazol (a standard veterinary anesthetic mixture of tiletamine and zolazepam), or 10 mg/kg of Ketaset (ketamine hydrochloride), rendering them unconscious for approximately 1/2 h. They were weighed on a Mettler Toledo SM 30-K digital electronic balance to an accuracy of ± 10 g, and their crown–rump length was measured to centimeter accuracy. Each animal was laid upon its side on the radiographic table, and midsagittal radiographs (hereafter x rays) were made of the head and neck region (1/40-s exposure time, 100 mA, 92–116 kV depending on animal size). To allow accurate determination of absolute size from the x-ray films, a 3-cm lead reference strip was placed within the area of exposure, held above the table by a cardboard support at the animals' midsagittal height.

X-ray clarity was sufficient to delineate the outlines of both the oral and nasal vocal tracts, as well as the location of the glottis (Fig. 1). The location of the glottis was made clearer by the presence of a subhyoid air sac in this species (Geist, 1933), which opens into the larynx via a thin tube directly above the level of the ventricular folds, and was clearly visible in the x-ray images.

Vocal tract length (VTL) was determined from tracings of the x-ray images using a Wacom ArtZ UD 0608 digitizing tablet and NIH Image software (version 1.52, available free on the Internet from NIH). A curvilinear line equidistant from the medial and external walls of the oral or nasal vocal tract was drawn from the middle of the glottis to the opening of the lips or external nares, respectively (Fig. 1). This measure of VTL is acoustically motivated: It follows the path of plane-wave propagation of sound from the glottis to the oral or nasal radiation site, and should on theoretical grounds be associated with formant frequency dispersion (Fant, 1960). The VTL, in pixels, was recorded along with the length of the 3-cm lead strip, in pixels, which was later used to derive the actual VTL in cm. The length of the entire skull from the occipital ridge to the front of the incisors was also recorded.

The accuracy of this technique was very high: ten repeated measures of one animal's x ray yielded standard deviations of 5 pixels or less (1 mm or less). Standard errors were a fraction of a millimeter for all x-ray measures.

C. Acoustic measures

Recordings were made with a Sony WM-D3 Walkman Professional recorder and Sony PC-62 microphone using Maxell XL II high-bias cassette tapes. Animals were recorded in their home cages before being captured or anesthetized. Lip configuration was carefully observed during recording, because macaques, like humans, can substantially shorten their vocal tract by retracting the lips, or lengthen it by protruding the lips (Hauser and Schön-Ybarra, 1994). Animals typically emitted aggressive ''pant-threats'' and ''pant-barks'' (Hauser *et al.*, 1993) when faced or stared at by an unfamiliar observer. These vocalizations were made with an almost completely closed mouth and no lip protrusion (the ''tense-mouth face'' of van Hoof, 1967), and thus should be closely registered with the anatomical measure of oral vocal tract length employed in this study. Threat calls are brief, noisy, coughlike calls (Fig. 2 and Bercovitch *et al.*, 1995), with a very broadband source, making them well-suited for formant frequency measurement (unlike a high-frequency tonal call, where formants are only detectable if a harmonic and formant coincide, see Lieberman and Blumstein, 1988; Ryalls and Lieberman, 1982). Therefore although some other call types were recorded (mainly coos), all acoustic analyses were performed on these threat vocalizations.

Recordings were digitized at 16 bits quantization and 22-kHz sampling rate using the built-in sound input on an Apple Power Macintosh 6100/60, using Macromedia SoundEdit 16 software. Formant frequencies were measured using LPC analysis (Markel and Gray, 1976; Wakita, 1976), implemented via autocorrelation in Matlab 4.2 (using the Matlab Signal Processing Toolbox). LPC analysis has been

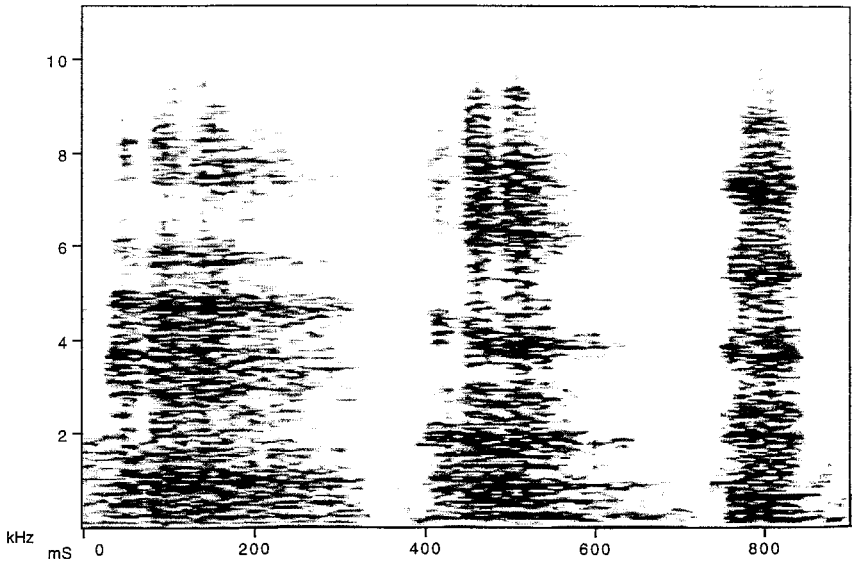


FIG. 2. Spectrograms of threat vocalizations from three different individual *M. mulatta* of increasing size. Note the decreasing formant dispersion (from left to right: 2.6 kg female; 5.3 kg female; 9.2 kg female) (2048-point Hamming window FFT, 50% frame overlap, 176 Hz bandwidth).

used to analyze similar primate vocalizations (Owren and Bernacki, 1988; Rendall, 1996), and its applicability has been verified by Owren (1990a, b).

Vocal tract length variation between monkeys should theoretically result in a variable number of formants below the Nyquist frequency, with short vocal tracts producing few formants and long tracts producing many. Using a simple quarter-resonator tube model (Lieberman and Blumstein, 1988), a 5-cm vocal tract would have three formants below our Nyquist frequency of 11 kHz, while a 10-cm vocal tract would have seven. A variable number of poles was used in the LPC analysis (3–9; LPC order=8–20). All LPC measurements were visually verified by superimposing the LPC-derived frequency response over an 512-point fast Fourier transform (FFT) of the same time slice, allowing the user to empirically determine the optimum number of poles for each call by trial and error.

The 512-sample arrays from locations chosen interactively by the user were input to the LPC function (no preemphasis, no weighting window). At least five LPC spectra were derived by analyzing successive time-slices of each call. Due to the noisy and random nature of the glottal source in pant-threats, all formants were not excited in each time slice; therefore, the successive LPC spectra were averaged together resulting in a single long-term averaged LPC spectrum. Typically, formant locations did not change appreciably during the course of a call; occasionally there was slight ($\pm$100 Hz) variation in a given formant frequency over the course of the call (probably due to slight mouth closing or opening).

Animals varied greatly in the number of usable vocalizations produced, with some producing only a few and others producing dozens. However, due to noise in the animal room from cage movements and other animals, most call recordings obtained were unanalyzable. Fortunately, formant frequency measurements are extremely consistent from one call token to the next (see Results). Thus for most analyses, the single best call (highest signal-to-noise ratio) was chosen (if there were multiple possibilities, the first one on the tape was chosen). A single call was used for each animal because for many animals only a single good-quality call was available, and this avoided the statistical complications caused by mixing single measurements for some animals with mean values for others.

Formant dispersion (D_f), which is the average distance between each adjacent pair of formants, was calculated using the following formula:

$$D_f = \frac{\Sigma_{i=1}^{N-1} F_{i+1} - F_i}{N-1}, \tag{1}$$

where D_f is the formant dispersion (in Hz), N is the total number of formants measured, and F_i is the frequency (in Hz) of formant i.

II. RESULTS

Summary data for each variable measured are given in Table I, and a summary of all the correlations examined in this study is given in Table II. In general, intercorrelations were very strong between vocal tract length (VTL), formant dispersion, and both measures of body size. All of the variables measured were roughly normally distributed, so the use of parametric statistics (i.e., regression analysis) was justified.

Although males tended to be slightly larger than females, there was no significant sex difference in any anatomical or acoustic variable measured (unpaired t tests, $p>0.05$). Therefore, unless otherwise noted all of the analyses reported here combine data from males and females.

A. Anatomical correlations

Extremely tight correlations were found among measurements of body size, skull length, and vocal tract length. Body weight and body length were highly correlated

TABLE I. Basic descriptive data for acoustic and anatomical variables. (σ is the standard deviation, S.E. the standard error of the mean, N the count, and ''min'' and ''max'' the minimum and maximum values).

	Mean	σ	S.E.	N	Min.	Max.	# Missing
Age (yr)	4.667	2.590	0.610	18	1.000	9.000	2
Weight (kg)	7.128	3.774	0.844	20	2.600	15.580	0
Body length (cm)	47.417	6.576	1.550	18	35.000	60.000	2
Skull length (cm)	12.204	1.558	0.348	20	9.254	15.022	0
Oral VTL (cm)	7.850	1.290	0.289	20	5.514	9.739	0
Nasal VTL (cm)	8.915	1.348	0.301	20	6.739	10.891	0
# Formants	4.800	1.399	0.313	20	3.000	7.000	0
Formant D (kHz)	2.146	0.700	0.157	20	1.250	3.510	0
Log_{10}wt	0.792	0.243	0.054	20	0.415	1.193	0

($r=0.911$, $p<0.0001$), as were body length and $\log_{10}$ body weight ($r=0.953$, $p<0.0001$) (Fig. 3). Because body weight should theoretically be proportional to the cube of a linear dimension, $\log_{10}$ body weight is used hereafter (logarithms were not taken for all measurements because the relationships between body length, skull length, VTL, and formant dispersion appear to be linear).

Skull length was correlated with body length ($r=0.974$, $p<0.0001$) and $\log_{10}$ weight ($r=0.955$, $p<0.0001$). VTL was correlated with skull length (oral VTL, $r=0.957$, $p<0.0001$; nasal VTL, $r=0.927$, $p<0.0001$), as was expected given that the main dimensions of the vocal tract are delineated by the skull (Fig. 3). As expected due to the intercorrelations between body size, skull length, and VTL, VTL correlated with $\log_{10}$ body weight (oral VTL, $r=0.947$, $p<0.0001$; nasal VTL, $r=0.906$, $p<0.0001$) and body length (oral VTL, $r=0.950$, $p<0.0001$, Fig. 3; nasal VTL, $r=0.928$, $p<0.0001$).

B. Acoustic correlations

The number of usable pant-threat calls obtained varied greatly among animals (range 1–10, mode=1, for four animals). For four animals, seven or more calls were obtained, allowing an analysis of consistency in formant locations among the calls of a particular individual. The formant values for these monkeys are shown in Fig. 4, which illustrates the impressive consistency of formant locations for each individual, and the substantial differences between individuals. Therefore, the rest of the analyses were performed with only one call per animal, as explained in the Methods section.

There was variation in the number of formants found for different animals. The number of formants below the 11 kHz Nyquist frequency (''formant density'') varied from three to seven (mean 4.8), and as predicted by theory was correlated with oral vocal tract length ($r=0.828$, $p<0.0001$)and inversely correlated with formant dispersion ($r=-0.918$, $p<0.0001$). Because formant density is an integer value, it is less sensitive than formant dispersion and is not further used (see Sec. III).

There were always at least three formants for any given call. Therefore, acoustic analyses were done in two ways for each call: including all formants present (thus making use of all information available for any given call), or including only the lowest three formants per call (assuring equality of analysis between animals). The pattern of results was identical for the two analyses, and therefore only the measure of formant dispersion defined in Eq. (1), which includes all available formats, is reported.

Formant dispersion was correlated significantly and negatively with VTL, as predicted by the source–filter theory of vocal production (oral VTL, $r=-0.915$, $p<0.0001$; nasal VTL, $r=-0.852$, $p=0.0001$). Given the anatomical correlations described above, it is thus unsurprising that formant dispersion correlated with $\log_{10}$ body weight ($r=-0.868$, $p<0.0001$) and body length ($r=-0.816$, $p<0.0001$). Stepwise multiple regression analysis indicated that $\log_{10}$ body weight accounted for the most variance in formant dispersion (partial correlation coefficients: age -0.713, body length -0.821, $\log_{10}$ body weight -0.876), and that it alone accounted for all significant variance in formant dispersion.

The formant dispersion (D_{pred}, in Hz) predicted by a simple one-tube model of the vocal tract (without end correction, since no data was available on size of the oral opening) is

TABLE II. Correlation coefficients between the various acoustic and anatomical variables measured in this study. All correlations are significant at the $p<0.0001$ level.

	1	2	3	4	5	6	7
1. Log_{10} wt	1.000	0.942	0.944	0.904	−0.723	−0.886	0.953
2. Skull length	0.942	1.000	0.963	0.939	−0.711	−0.869	0.974
3. Oral VTL	0.944	0.963	1.000	0.972	−0.724	−0.922	0.950
4. Nasal VTL	0.904	0.939	0.972	1.000	−0.625	−0.868	0.928
5. $F1$	−0.723	−0.711	−0.724	−0.625	1.000	0.699	−0.702
6. Formant D	−0.886	−0.869	−0.922	−0.868	0.699	1.000	−0.816
7. Body length	0.953	0.974	0.950	0.928	−0.702	−0.816	1.000

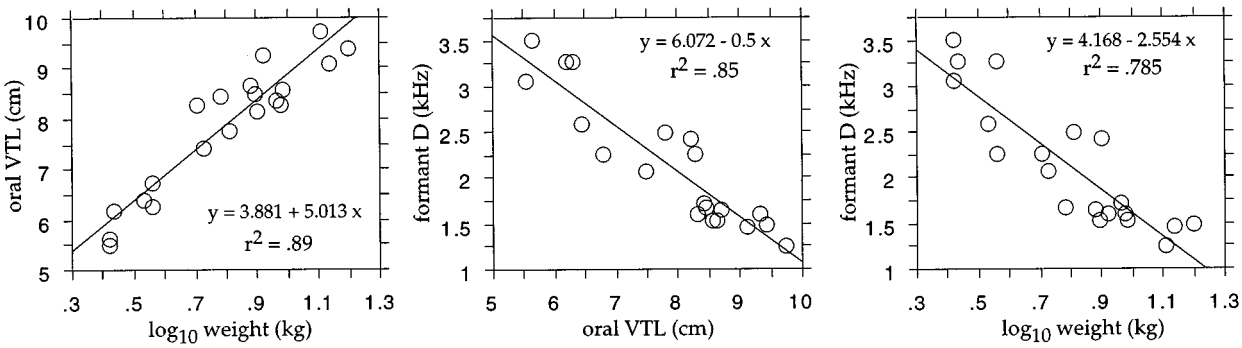


FIG. 3. Bivariate plots illustrating intercorrelations of the base 10 logarithm of weight (kg), oral vocal tract length (cm), and formant dispersion (kHz).

$$D_{\text{pred}} = \frac{c}{2L}, \tag{2}$$

where c is the speed of sound (335 m/s) and L is the vocal tract length in m. This prediction is the same for open or closed tubes (where the resonances are related as 2,4,6,...), or half-open tubes (where the relation is 1,3,5,...). The formant dispersion observed did not differ significantly from that predicted by Eq. (2) (paired t tests, $t=0.63$, $p=0.53$). Thus although the x-ray data indicate that the monkey vocal tract is not a tube of exactly uniform diameter, use of a more complex model of vocal tract anatomy than a simple tube (as suggested by Shipley *et al.*, 1991) does not seem warranted by the current data. The good fit between predicted and actual values also provides further evidence of the usefulness of low-order LPC for isolating and measuring spectral peaks of nonhuman vocal tracts (Owren and Bernacki, 1988; Owren, 1990a,b). Because the excitation signal for the threat vocalizations analyzed here is coughlike and noisy (presumably generated by noise at the glottis), it is plausible to interpret the low-order LPC spectrum as characterizing the vocal tract filter function, and thus the spectral peaks in the LPC spectrum as formant frequencies.

These results were consistent across both ages and sexes. When only adult animals were included in the analysis (i.e., animals aged 4 years or older), the correlation between $\log_{10}$ body weight and formant dispersion dropped (due to the smaller sample size) but was still very significant statistically ($N=11$, $r=-0.73$, $p=0.008$). The correlation also held for juveniles only ($N=7$, $r=-0.84$, $p=0.015$). Similarly, the correlation between $\log_{10}$ body weight and formant dispersion was strong in the 11 males considered separately ($r=-0.896$, $p=0.0002$), as in the nine females separately ($r=-0.913$, $p=0.0006$). Thus the relationship between body size and acoustic output described here does not result simply from differences between adults and juveniles, or between males and females. Instead, it appears to result from a direct anatomical relationship between body size, vocal tract length, and formant dispersion.

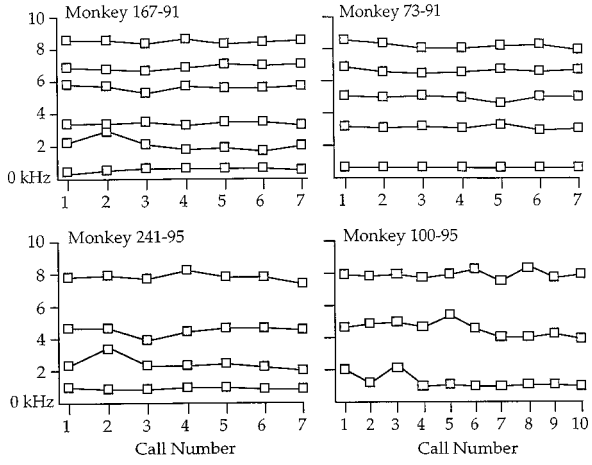


FIG. 4. Consistency of formant frequency measurements across different calls for each of four individual monkeys.

III. DISCUSSION

The results reported here indicate that formant dispersion, which is determined by vocal tract length (VTL), is closely correlated with body weight and length in rhesus macaques. As predicted by source–filter acoustic theory, in tandem with the tight anatomical correlations reported above, the spacing between formants provides an accurate means of predicting a vocalizer's body mass and/or length. Formant dispersion in the threat vocalizations of rhesus macaques thus provides an excellent *potential* cue to body size in the population of rhesus macaques studied here (r^2 greater than 75%). Whether rhesus macaques use this information remains unknown, but a growing body of data (reviewed in the Introduction) strongly suggest that nonhuman animals attend to formant cues and make use of them in their species-specific communication systems (Dewson, 1964; Lieberman, 1968; Baru, 1975; Andrew, 1976; Richman, 1976; Heinz *et al.*, 1981; Seyfarth and Cheney, 1984; Heinz and Brady, 1988; Owren and Bernacki, 1988; Owren, 1990a,b; Sommers *et al.*, 1992; Hauser and Schön-Ybarra, 1994; Hauser *et al.*, 1993; Fitch, 1994; Fitch and Hauser, 1995; Rendall; 1996; Owren *et al.*, 1997). Rendall (1996) suggested that one use of formant information in macaque communication is to provide a relatively stable indicator of individual identity. The data in this paper support this idea (because formants were found to be very stable over multiple calls), and suggest that another potential use for formant information is to evaluate a vocalizer's body size.

A. Reliability of formant dispersion

Despite the common conception that F_0 provides a cue to body size (e.g., Morton, 1977), the preponderance of evidence now suggests that F_0 does not provide an accurate indication of size in adult humans (Lass and Brown, 1978; Künzel, 1989; Cohen *et al.*, 1980; van Dommellen, 1993) or in several other animals (McComb, 1991; Asquith and Altig, 1990; Sullivan, 1984). I suggest that this lack of correlation is due to the relative independence between body size and larynx size (which determines vocal fold length and hence lowest F_0). The cartilages of the larynx can grow independently of the rest of the body, and experience a hormone-dependent growth spurt in many mammals (Tuohimaa *et al.*, 1981; Beckford *et al.*, 1985; Hollien, 1960) that leads to a decoupling of body size and F_0, particularly in males. In contrast, vocal tract length is anatomically dependent upon skull size, which is in turn closely correlated with body size in all mammal species which have been examined (Morita and Ohtsuki, 1973; Dechow, 1983; Janis, 1990; Valkenburgh, 1990; Alcantara *et al.*, 1991; Sharma, 1990). The results of the current study indicate that (a) body size, skull size, and VTL are anatomically correlated in macaques, and (b) VTL determines formant dispersion, which thus provides an acoustic indication of body weight and mass. Hence, unlike F_0, formant dispersion should provide a robust, reliable cue to body size in most mammals and perhaps terrestrial vertebrates in general.

Two key factors underlie this reliability: the link between formant dispersion and VTL, and the link between VTL and body size. Each of these are discussed in turn. Because tract length affects the overall pattern of formant frequencies, any given formant provides some information regarding tract length. Thus, for example, the lowest formant ($F1$) could potentially be used as a cue to tract length. However, using one formant as the sole cue to VTL would entail several problems. First, if the frequency band of the chosen formant is obscured by environmental noise or degradation, no VTL information would be available. In contrast, formant dispersion capitalizes on the redundancy of the formant spacing pattern, and is thus robust to the degradation of information in any one (or even several) formants. Second, individual formants are more sensitive to changes in the vocal tract transfer function (as seen in human vowels and, to a lesser degree, in baboon grunts: Owren *et al.*, 1997). As a statistical measure encompassing all formant information, formant dispersion is less sensitive to deviations in a single formant.

Another potential cue to VTL is ''formant density'' (the number of formants in a particular frequency range, e.g., in this study the total number of formants below the 11-kHz Nyquist frequency), which is inversely related to formant dispersion. The main problem with density as an acoustic cue to tract length is that, as an integer measure, it is considerably less sensitive to changes in VTL than formant dispersion, which is as accurate as discrimination of individual formants [formant difference limens are 1%–5% in both humans and macaques, Sommers *et al.* (1992)]. Of course, it would be possible to devise more accurate formant density indices, but the simplicity and numerical tractability of formant dispersion recommend it. Thus formant dispersion is accurate, robust to environmental deterioration, and resistant to errors due to variability in individual formants.

B. Source of unreliability for vocal tract cues

When it comes to providing vocal tract information, not all calls are equal. If the source contains energy at only a few, widely spaced frequencies (e.g., a high-pitched tonal call), it will provide little information about the vocal tract transfer function. Thus low-pitched calls provide for better resolution of formants than do sounds with high F_0's (Ryalls and Lieberman, 1982). Better yet, a noisy or impulsive source is ideally suited for accurately outlining the transfer function (e.g., Tartter and Braun, 1994 with whispers). The accuracy of formant dispersion as an acoustic cue to body size will thus depend on the particular type of call, with harsh, noisy or impulsive calls (like the threat vocalizations studied here) being ideal, and high-pitched tonal calls being worst. As well as having clear methodological implications, this observation leads to an interesting behavioral prediction. In contexts where accurate information about size is favored (e.g., a large animal demonstrating its size to competitors or mates), noisy or low-pitched calls might be expected since they provide more accurate delineation of vocal tract cues (see also Peters, 1984). This provides alternative grounds for Morton's (1977) ''motivational-structural rule'' that low, noisy sounds accompany aggression.

Another potential source of unreliability is changes in vocal tract length due to articulatory movements. Opening or nearly closing the mouth, or (in animals with mobile lips) protruding or retracting the lips will alter tract length to some degree (around 20% in humans, Fant, 1960), as will raising or lowering the larynx. However, the manipulation of VTL via lip movements does not necessarily render it (or related acoustic cues) unreliable. There is still a maximum VTL, fixed by skull dimensions, which is attained when the mouth is nearly closed and the lips are pursed. This articulatory position is characteristic of the threat vocalizations studied here, along with threats in a wide variety of other species (see below). The observed correlation between lip rounding and threat raises an interesting possibility. If listeners associate long vocal tracts with larger bodies, an animal that elongates its vocal tract while vocalizing would maximize its perceived body size. This suggests that aggressive vocalizations made with nearly closed mouths and protruded lips would be more threatening, and submissive vocalizations made with lips retraced more appeasing, and that the evolutionary origin of certain facial expressions is not purely visual, but arises also from their acoustic effects (Ohala, 1980; Fridlund, 1994; Fitch, 1994; Fitch and Hauser, 1995).

Vocal tract elongation is associated with threat behavior in many species. van Hooff (1967; p. 18), in his general review of primate facial displays, cited the ''tense-mouth face'' in which the ''mouth corners are brought forward... as a result the mouth often looks like a narrow slit.'' This display is usually performed by a dominant animal immediately preceding an attack, and is associated with a low-pitched bark in at least some species (chimpanzees and baboons). Similar ''pouts'' or ''pucker'' behavior is associated with

threat in other primates (Kaufman, 1967; Fitch, 1994; Epple, 1967), seals (Miller, 1975), bears (Pruitt and Burghardt, 1977), and in dogs and other canids (Fox and Cohen, 1977).

Similarly, the use of mouth corner retraction to signal submission and absence of threat appears to be extremely widespread in mammals. A vast number of species, including dogs, seals, ungulates, and almost all primates share a submissive ''grin'' facial display (van Hooff, 1967, 1972; Miller; 1975; Fox and Cohen, 1977; Oppenheimer, 1977; Marler and Tenaza, 1977; Gautier and Gautier, 1977; Walther, 1977, 1984; Goodall, 1986; Preuschoft, 1992, 1995). This expression is variously ascribed a defensive or submissive role and is occasionally seen during play. The human smile appears to be an example of this display, although its use has been extended into nonaggressive situations to denote an unthreatening or friendly attitude (van Hooff, 1972; Preuschoft, 1992, 1995). These observations prompted Ohala (1980) to suggest that shortening the vocal tract in order to seem smaller would provide an acoustic (rather than visual) function for the human smile. The data in this paper are consistent with this hypothesis.

Finally, the anatomical correlation between VTL and skull size reported here and suspected in most mammals does not hold for all vertebrates. In birds, the sound-producing source is the syrinx, which lies at the base of the trachea near the lungs (Nowicki and Marler, 1988). Thus the trachea is an integral part of the bird vocal tract, and the length of the trachea must be added to the length of the oral/nasal cavities to derive the bird's total vocal tract length, which is not limited by the size of the skull. The trachea (like the larynx) is a flexible structure, floating free from any rigid skeletal attachments. It is thus interesting that a wide variety of birds exhibit tracheal elongation where the trachea is looped in great coils within the chest or sternum (Berndt, 1938; Neimeier, 1979). This bizarre character is found only in large birds, and often only in males, suggesting that it may represent an adaptation for acoustically exaggerating body size (Fitch, in preparation, 1994).

A second potential exception is anatomically modern *Homo sapiens*. Our species has abandoned the structural pattern typical of mammalian vocal tract anatomy. In humans the larynx has descended from its normal position at the back of the mouth to a position deeper in the throat (Negus, 1949; Laitman and Crelin, 1976; Leiberman, 1984; Flügel and Rohen, 1991), which allows the tongue body to move freely back and forth, thus creating a wider variety of vocal tract area functions than are possible in nonhumans (Lieberman, 1968; Lieberman *et al.*, 1969). Though typically viewed as an adaptation for articulate speech (Lieberman, 1975, 1984), this speech advantage would only be gained with considerable laryngeal descent (several cm). Presuming that this anatomical reconfiguration proceeded gradually over evolutionary time, what selective advantage was provided by the descent of the larynx before the conditions for improved speech were met?

The descent of the larynx from the standard mammalian position has the effect of elongating the vocal tract, possibly freeing human tract length from the skeletal size constraints described above for most mammals. Since vocal tract length is used by human listeners as a cue to body size (Fitch, 1994), it is possible that the original function of the descending larynx in early hominids was to exaggerate body size. This idea gains support from the observation that there is sexual dimorphism in the degree to which the human larynx descends (a full vertebra lower in males than females, Senecail, 1979; Harrison, 1995), with no accompanying advantage for males' vowel clarity over that of females. The degree to which vocal tract length, formant dispersion, and body size are related in humans remains an interesting matter for further research.

C. Vocal tract cues and human speech perception

One key phenomenon in human speech perception that may be related to the data in this paper is vocal tract normalization, which is the use of overall formant pattern to ''normalize'' the vowels of a given speaker (Ladefoged and Broadbent, 1957; Nearey, 1978; Lieberman, 1984). Because of variations in vocal tract length, the $F1$-$F2$ vowel space of a child is quite different from that of an adult (Peterson and Barney, 1952). A result is that children imitating adults, or adults listening to children, need to adjust their perception of a speaker's vowels to their VTL-related formant space. If a mechanism for estimating body size from formant dispersion existed in our prelinguistic ancestors, this could have provided a preadaptive basis for vocal tract normalization in humans.

The current results joins a collection of recent studies suggesting that the role of the supralaryngeal vocal tract in nonhuman animal communication may be more prominent than traditionally realized (e.g., Bauer, 1987; Owren, 1990a,b; Owren *et al.*, 1997; Shipley *et al.*, 1991; Hauser and Schön-Ybarra, 1994; Owren and Bernacki, 1988; Rendall, 1996). Although formant frequency perception plays a fundamental role in human speech perception (Joos, 1948; Lieberman and Blumstein, 1988), the evolutionary basis for this is poorly understood, mainly because we have so little information about the uses to which formant frequencies may be put in animal communication systems. The results of this work suggest that the use of formant frequency patterns to gauge body size could have been a factor promoting the original evolution of formant perception, a capability later put to such extensive and sophisticated use in human language.

ACKNOWLEDGMENTS

Many thanks are due to David Lee-Parritz for the use of the facilities at the New England Primate Center. I am very grateful to Dr. Lee-Parritz, Barbara Brewer, Virginia Rainville, and Nirah Shomer for their invaluable help in gathering data. Miguel Schön-Ybarra provided assistance in interpreting the x-ray data. Comments by Marc Hauser, Caroline Henton, Nelson Kiang, Daniel Lieberman, Philip Lieberman, Michael Owren, Ken Stevens, and an anonymous reviewer are gratefully acknowledged. This work was supported by an NIH Postdoctoral Fellowship to the author.

Alcantara, M., Diaz, M., and Pulido, F. J. P. (**1991**). ''Variabilidad en las relaciones alométricas entre el peso y las medidas craneales en el raton de campo *Apodemus sylvaticus* L. Efectos sobre su utilidad en estudios de

ecologia trofica de aves rapaces,'' Doñana, Acta Vertebrata **18**, 205–216.
Alexander, R. M. (**1996**). ''Biophysical problems of small size in vertebrates,'' Symp. Zoological Soc. London **69**, 3–14.
Andrew, R. J. (**1976**). ''Use of formants in the grunts of baboons and other nonhuman primates,'' Ann. (N.Y.) Acad. Sci. **280,** 673–693.
Asquith, A., and Altig, R. (**1990**). ''Male call frequency as a criterion for female choice in *Hyla cinerea*,'' J. Herpetology **24**, 198–201.
Baru, A. V. (**1975**). ''Discrimination of synthesized vowels [a] and [i] with varying parameters (fundamental frequency, intensity, duration and number of formants) in dog,'' *Auditory Analysis and Perception of Speech*, edited by G. Fant and M. A. A. Tatham (Academic, New York).
Bauer, H. R. (**1987**). ''Frequency code: orofacial correlates of fundamental frequency,'' Phonetica **44**, 173–191.
Beckford, N. S., Rood, S. R., and Schaid, D. (**1985**). ''Androgen stimulation and laryngeal development,'' Ann. Otol. Rhinol. Laryngol. **94,** 634–640.
Bercovitch, F. B., Hauser, M. D., and Jones, J. H. (**1995**). ''The endocrine stress response and alarm vocalizations in rhesus macaques,'' Animal Beh. **49,** 1703–1706.
Berndt, R. (**1938**). ''Intrasternale Trachealschlingen bei Vogëln,'' Gegenbauer's Morphologische Jahrbuch **82**, 27–117.
Clutton-Brock, T. H., and Albon, S. D. (**1979**). ''The roaring of red deer and the evolution of honest advertising,'' Behaviour **69**, 145–170.
Clutton-Brock, T. H., Harvey, P. H., and Rudder, B. (**1977**). ''Sexual dimorphism, socionomic sex ratio and body weight in primates,'' Nature (London) **269,** 797–800.
Cohen, J. R., Crystal, T. H., House, A. S., and Neuburg, E. P. (**1980**). ''Weighty voices and shaky evidence: a critique,'' J. Acoust. Soc. Am. **68,** 1884–1886.
Darwin, C. (**1871**). *The Descent of Man and Selection in Relation to Sex* (Murray, London).
Davies, N. B., and Halliday, T. R. (**1978**). ''Deep croaks and fighting assessment in toads,'' Nature (London) **274,** 683–685.
Dechow, P. C. (**1983**). ''Estimation of body weights from craniometric variables in baboons,'' Am. J. Phys. Anthropol. **60**, 113–123.
Dewson, J. H. (**1964**). ''Speech sound discrimination by cats,'' Science **141,** 555–556.
Epple, G. (**1967**). ''Vergleichende Untersuchungen über Sexual- und Sozialverhalten der Krallenaffen (Hapalidae),'' Folia Primatogica **7**, 37–65.
Fant, G. (**1960**). *Acoustic Theory of Speech Production* (Mouton, The Hague).
Fitch, W. T. (**1994**). ''Vocal tract length perception and the evolution of language,'' Ph.D. thesis, Brown University.
Fitch, W. T., and Hauser, M. D. (**1995**). ''Vocal production in nonhuman primates: Acoustics, physiology and functional constraints on 'honest' advertising,'' Am. J. Primatology **37**, 191–219.
Fitch, W. T. (**in preparation**). ''Acoustic exaggeration of size in birds by tracheal elongation: comparative and theoretical analyses.''
Flügel, C., and Rohen, J. W. (**1991**). ''The craniofacial proportions and laryngeal position in monkeys and man of different ages. (A morphometric study based on CT-scans and radiographs),'' Mech. Aging Develop. **61**, 65–83.
Fox, M. W., and Cohen, J. A. (**1977**). ''Canid communication,'' in *How Animals Communicate*, edited by T. A. Sebeok (Indiana U. P., Bloomington, IN).
Fridlund, A. J. (**1994**). *Human Facial Expression: An Evolutionary View* (Academic, San Diego).
Gautier, J. P., and Gautier, A. (**1977**). ''Communication in Old World monkeys,'' in *How Animals Communicate*, edited by T. A. Sebeok (Indiana U. P., Bloomington, IN).
Geist, F. D. (**1933**). ''Nasal cavity, larynx, mouth, and pharynx,'' in *The Anatomy of the Rhesus Monkey*, edited by Hartman and Straus (Hafner, New York).
Goldstein, U. G. (**1980**). ''An articulatory model for the vocal tracts of growing children,'' D. Sc. thesis, Massachusetts Institute of Technology.
Goodall, J. (**1986**). *The Chimpanzees of Gombe: Patterns of Behavior* (Harvard U. P., Cambridge, MA).
Harrison, D. F. N. (**1995**). *The Anatomy and Physiology of the Mammalian Larynx* (Cambridge U. P., Cambridge, New York).
Harvey, P. H. (**1990**). ''Life-history variation: Size and mortality patterns,'' in *Primate Life History & Evolution*, edited by C. J. DeRousseau (Wiley–Liss, New York).
Hauser, M. D. (**1993**). ''The evolution of nonhuman primate vocalizations: effects of phylogeny, body weight and social context,'' Am. Nat. **142,** 528–542.
Hauser, M. D., Evans, C. S., and Marler, P. (**1993**). ''The role of articulation in the production of rhesus monkey (*Macaca mulatta*) vocalizations,'' Animal Beh. **45,** 423–433.
Hauser, M. D., and Schön-Ybarra, M. (**1994**). ''The role of lip configuration in monkey vocalizations: Experiments using xylocaine as a nerve block,'' Brain Lang. **46**, 232–244.
Heinz, R. D., and Brady, J. V. (**1988**). ''The acquisition of vowel discriminations by nonhuman primates,'' J. Acoust. Soc. Am. **84,** 186–194.
Heinz, R. D., Sachs, M. B., and Sinnott, J. M. (**1981**). ''Discrimination of steady-state vowels by blackbirds and pigeons,'' J. Acoust. Soc. Am. **70,** 699–706.
Hollien, H. (**1960**). ''Some laryngeal correlates of vocal pitch,'' J. Speech Hear. Res. **3,** 52–58.
Janis, C. (**1990**). ''Correlation and cranial and dental variables with body size in ungulates and macropodoids,'' in *Body Size in Mammalian Paleobiology: Estimation and Biological Implications*, edited by J. Damuth and B. J. MacFadden (Cambridge U. P., Cambridge, England).
Joos, A. M. (**1948**). ''Acoustic phonetics,'' Language **24** (Suppl.), 1–136.
Kaufmann, J. H. (**1967**). ''Social relations of adult males in a free-ranging band of rhesus monkeys,'' in *Social Communication Among Primates*, edited by S. A. Altmann (University of Chicago, Chicago).
Kingdon, J. (**1974**). *East African Mammals: Vol. II Part A: Insectivores and Bats* (Academic, New York).
Künzel, H. J. (**1989**). ''How well does average fundamental frequency correlate with speaker height and weight?,'' Phonetica **46**, 117–125.
Ladefoged, P., and Broadbent, D. E. (**1957**). ''Information conveyed by vowels,'' J. Acoust. Soc. Am. **39,** 98–104.
Laitman, J. T., and Crelin, E. S. (**1976**). ''Postnatal development of the basicranium and vocal tract region in man,'' in *Symposium on Development of the Basicranium*, edited by J. F. Bosma (U. S. Government Printing Office, Washington, DC).
Lass, N. J., and Brown, W. S. (**1978**). ''Correlational study of speakers' heights, weights, body surface areas, and speaking fundamental frequencies,'' J. Acoust. Soc. Am. **63,** 1218–1220.
Lieberman, P. (**1968**). ''Primate vocalization and human linguistic ability,'' J. Acoust. Soc. Am. **44,** 1574–1584.
Lieberman, P. (**1975**). *On the Origins of Language* (Macmillan, New York).
Lieberman, P. (**1984**). *The Biology and Evolution of Language* (Harvard U. P., Cambridge, MA).
Lieberman, P., and Blumstein, S. E. (**1988**). *Speech Physiology, Speech Perception, and Acoustic Phonetics* (Cambridge U. P., New York).
Lieberman, P. H., Klatt, D. H., and Wilson, W. H. (**1969**). ''Vocal tract limitations on the vowel repertoires of rhesus monkey and other nonhuman primates,'' Science **164,** 1185–1187.
Markel, J. D., and Gray, A. H. (**1976**). *Linear Prediction of Speech* (Springer-Verlag, New York).
Marler, P., and Tenaza, R. (**1977**). ''Signaling behavior of apes with special reference vocalization,'' in *How Animals Communicate*, edited by T. A. Sebeok (Indiana U. P., Bloomington, IN).
Martin, W. F. (**1972**). ''Evolution of vocalizations in the genus *Bufo*,'' in *Evolution in the genus Bufo*, edited by W. F. Blair (University of Texas, Austin, TX), pp. 279–309.
McComb, K. E. (**1991**). ''Female choice for high roaring rates in red deer, *Cervus elaphus*,'' Animal Beh. **41,** 79–88.
Miller, E. H. (**1975**). ''A comparative study of facial expressions of two species of pinnipeds,'' Behaviour **53**, 268–284.
Morita, S., and Ohtsuki, F. (**1973**). ''Secular changes of the main head dimensions in Japanese,'' Human Biol. **45**, 151–165.
Morton, E. S. (**1977**). ''On the occurrence and significance of motivation-structural rules in some bird and mammal sounds,'' Am. Nat. **111,** 855–869.
Nearey, T. (**1978**). *Phonetic Features for Vowels* (Indiana University Linguistics Club, Bloomington).
Negus, V. E. (**1949**). *The Comparative Anatomy and Physiology of the Larynx* (Hafner, New York).
Niemeier, M. M. (**1979**). ''Structural and functional aspects of vocal ontogeny in *Grus canadensis* (*Gruidae: Aves*),'' Ph.D. thesis, University of Nebraska, Lincoln.
Nowicki, S., and Marler, P. (**1988**). ''How do birds sing?,'' Music Percept. **5**, 391 426.
Ohala, J. J. (**1980**). ''ABSTRACT: The acoustic origin of the smile,'' J. Acoust. Soc. Am. Suppl. 1 **68,** S33.
Oppenheimer, J. R. (**1977**). ''Communication in New World monkeys,'' in

How Animals Communicate, edited by T. A. Sebeok (Indiana U. P., Bloomington, IN).
Owren, M. J. (**1990a**). ''Acoustic classification of alarm calls by vervet monkeys (*Cercopithecus aethiops*) and humans: I. Natural calls,'' J. Comp. Psych. **104,** 20–28.
Owren, M. J. (**1990b**). ''Acoustic classification of alarm calls by vervet monkeys (*Ceropithecus aethiops*) and humans: II. Synthetic calls,'' J. Comp. Psych. **104,** 29–40.
Owren, M. J., and Bernacki, R. (**1988**). ''The acoustic features of vervet monkey (*Ceropithecus aethiops*) alarm calls,'' J. Acoust. Soc. Am. **83,** 1927–1935.
Owren, M. J., Seyfarth, R. M., and Cheney, D. L. (**1997**). ''The acoustic features of vowel-like *grunt* calls in chacma baboons (*Papio cyncephalus ursinus*): Implications for production processes and functions,'' J. Acoust. Soc. Am. **101**, 2951–2963.
Parker, G. A. (**1974**). ''Assessment strategy and the evolution of fighting behavior,'' J. Theor. Biol. **47,** 223–243.
Peters, G. (**1984**). ''On the structure of friendly close range vocalizations in terrestrial carnivores (Mammalia: Carnivora: Fissipedia),'' Z. Säugetierkunde **49**, 157–182.
Peters, R. H. (**1983**). *The Ecological Implications of Body Size* (Cambridge U. P., New York).
Peterson, G. E., and Barney, H. L. (**1952**). ''Control methods used in a study of vowels,'' J. Acoust. Soc. Am. **24,** 175–184.
Preuschoft, S. (**1992**). '' 'Laughter' and 'smiling' in Barbary macaques (*macaca sylvanus*),'' Ethology **91**, 220–236.
Preuschoft, S. (**1995**). '' 'Laughter' and 'smiling' in macaques: An evolutionary perspective,'' Ph.D. thesis, Utrecht University.
Pruit, C. H., and Burghardt, G. M. (**1977**). ''Communication in terrestrial carnivores: Mustelidae, Procyonidae and Ursidae,'' in *How Animals Communicate*, edited by T. A. Sebeok (Indiana U. P., Bloomington, IN).
Rendall, C. A. (**1996**). ''Social communication and vocal recognition in free-ranging rhesus monkeys (*Macaca mulatta*), Ph.D. thesis, University of California, Davis.
Richman, B. (**1976**). ''Some vocal distinctive features used by gelada monkeys,'' J. Acoust. Soc. Am. **60,** 718–724.
Ryalls, J. H., and Lieberman, P. (**1982**). ''Fundamental frequency and vowel perception,'' J. Acoust. Soc. Am. **72,** 1631–1634.
Ryan, M. J. (**1988**). ''Constraints and patterns in the evolution of anuran acoustic communication,'' in *The Evolution of the Amphibian Auditory System*, edited by B. Fritzch, M. J. Ryan, W. Wilczynski, T. E. Hetherington, and W. Walkowiak (Wiley, New York).
Schmidt-Nielsen, K. (**1984**). *Scaling: Why is animal size so important?* (Cambridge U. P., New York).
Schneider, R., Kuhn, H.-J., and Kelemen, G. (**1967**). ''Der Larynx des Männlichen *Hypsignathus monstrosus* Allen, 1861 (Pteropodidae, Megachiroptera, Mammalia),'' Z. wissenschaft. Zool. **175**, 1–53.
Schön, M. (**1971**). ''The anatomy of the resonating mechanism in howling monkeys,'' Folia Primatologica **15**, 117–132.
Senecail, B. (**1979**). *L'Os Hyoide: Introduction Anatomique a l'Etude de certains Mecanismes de la Phonation* (Faculté de Médicine de Paris, Paris).
Seyfarth, R. M., and Cheney, D. L. (**1984**). ''The acoustic features of vervet monkey grunts,'' J. Acoust. Soc. Am. **75,** 129–134.
Sharma, N. (**1990**). *Facial Growth and Body Size* (Agam Prakashan, Delhi).
Shipley, C., Carterette, E. C., and Buchwald, J. S. (**1991**). ''The effects of articulation on the acoustical structure of feline vocalizations,'' J. Acoust. Soc. Am. **89,** 902–909.
Sommers, M. S., Moody, D. B., Prosen, C. A., and Stebbins, W. C. (**1992**). ''Formant frequency discrimination by Japanese macaques (*Macaca fuscata*),'' J. Acoust. Soc. Am. **91,** 3499–3510.
Sullivan, B. K. (**1984**). ''Advertisement call variation and observations on breeding behavior of *Bufo debilis* and *B. punctatus*,'' J. Herpetology **18**, 406–411.
Tartter, V. C., and Braun, D. (**1994**). ''Hearing smiles and frowns in normal and whisper registers,'' J. Acoust. Soc. Am. **96,** 2101–2107.
Tuohimaa, P. T., Kallio, S., and Heinijoki, J. (**1981**). ''Androgen receptors in laryngeal cancer,'' Acta Oto-Laryngol. **91,** 159–164.
van Valkenburgh, B. (**1990**). ''Skeletal and dental predictors of body mass in carnivores,'' in *Body Size in Mammalian Paleobiology: Estimation and Biological Implications*, edited by J. Damuth and B. J. MacFadden (Cambridge U. P., Cambridge, England).
van Dommelen, W. A. (**1993**). ''Speaker height and weight identification: a reevaluation of some old data,'' J. Phonetics **21**, 337–341.
van Hooff, J. A. R. A. M. (**1967**). ''The facial displays of the Catarrhine monkeys and apes,'' in *Primate Ethology*, edited by D. Morris (Weidenfeld and Nicolson, London).
van Hooff, J. A. R. A. M. (**1972**). ''A comparative approach to the phylogeny of laughter and smiling,'' in *Nonverbal Communication*, edited by R. Hinde (Cambridge U. P., New York).
Wakita, H. (**1976**). ''Instrumentation for the study of speech acoustics,'' in *Contemporary Issues in Experimental Phonetics*, edited by N. J. Lass (Academic, New York).
Walther, F. R. (**1977**). ''Artiodactyla,'' in *How Animals Communicate*, edited by T. A. Sebeok (Indiana U. P., Bloomington, IN).
Walther, F. R. (**1984**). *Communication and Expression in Hoofed Mammals* (Indiana U.P., Bloomington, IN).

Sound radiation of 3-MHz driven gas bubbles

Siegfried Grossmann, Sascha Hilgenfeldt, and Detlef Lohse
Fachbereich Physik der Universität Marburg, Renthof 6, 35032 Marburg, Germany

Michael Zomack
Schering AG, Clinical Development, Müllerstr. 178, 13342 Berlin, Germany

(Received 29 October 1996; accepted for publication 14 April 1997)

The sound radiation of 3-MHz acoustically driven air bubbles in liquid is analyzed with respect to possible applications in second harmonic ultrasound diagnostics devices, which have recently come into clinical use. In the forcing pressure amplitude $P_a = 1-10$ atm and ambient radius $R_0 = 0.5-5$ μm parameter domain, a narrow regime around the resonance radius $R_0 \sim 1-1.5$ μm and relatively modest $P_a \sim 2-2.5$ atm is identified in which optimal sound yield in the second harmonic is achieved while maintaining spherical stability of the bubble. For smaller P_a and larger R_0 hardly any sound is radiated; for larger P_a bubbles become unstable toward nonspherical shape oscillations of their surface. The computation of these instabilities is essential for the evaluation of the optimal parameter regime. A region of slightly smaller R_0 and $P_a \sim 1-3$ atm is best suited to achieve large ratios of the second harmonic to the fundamental intensity. Spherical stability is guaranteed in the suggested regimes for liquids with an enhanced viscosity compared to water, such as blood. © *1997 Acoustical Society of America.* [S0001-4966(97)07307-4]

PACS numbers: 43.80.Qf, 43.35.Sx, 43.25.Yw [FD]

INTRODUCTION

Microbubbles, i.e., gas bubbles of a few μm diameter, have long been known to be very effective scatterers of ultrasound (cf., e.g., Ref. 1). Their scattering cross sections for MHz sound waves can be more than two orders of magnitude greater than their geometrical cross sections.[2] In the last decade, the concept of exploiting this property to perform refined ultrasound diagnostics with gas bubbles as echo contrast enhancers has enjoyed increasing attention.[3] The general idea of this technique is to inject a microbubble suspension into a vein and to study the blood flow by detecting the bubbles' sound echo reaction to an applied acoustical field. This leads to ultrasound images of higher contrast and quality as compared to conventional diagnostic techniques using only the ultrasound backscatter from the tissue itself.

The quality of an ultrasonogram mainly depends on its spatial resolution and its signal intensity; more specifically, on the ratio of signal intensities from ''desired'' echoes (such as the blood flow in bubble diagnostics) to ''background noise'' reflections (from surrounding tissue). Spatial resolution is, of course, limited by the wavelength of the ultrasound. But as the absorption of sound in tissue increases exponentially with frequency,[4,5] frequencies below 10 MHz are used in most clinical applications. As a typical value, we will choose an ultrasound driving frequency of $\omega_d/2\pi = 3$ MHz throughout this work.

In doing diagnostics with bubble suspensions, it is desirable to improve the signal to noise ratio. Namely, when detecting the emitted sound from the bubble at the driving frequency 3 MHz, the signal is obscured by the driving and its reflections from tissue. To improve the signal quality, it has recently been proposed[6,7] to detect *higher harmonics* of the driving frequency in the sound emission spectrum of the bubble. In view of the aforementioned strong damping of higher frequencies, the lowest (second) harmonic at 6 MHz is of particular interest. It can be selectively excited if the parameters are chosen appropriately. We expect that soft tissue, driven into the regime of nonlinear response by the strong driving, will also reflect part of the sound in higher harmonics. Also, the large amplitude driving signal itself will undergo nonlinear distortion, generating higher harmonic frequency components.[8] As in the case of the conventional method, it may therefore be necessary to focus on the *difference* between the reflected signal with and without injected microbubbles. However, by choosing the proper size of the bubbles and the proper forcing pressure amplitude, it is possible to enhance the bubbles' reflection signal in higher harmonics. The main focus of this paper is to suggest a parameter regime well suited for such an endeavor.

Experimental and theoretical research on bubble dynamics has received considerable attention since the discovery of single bubble sonoluminescence by Gaitan in 1990 (cf. Ref. 9) and the detailed experiments by the Putterman group at UCLA.[10–12] In those experiments single microbubbles are driven with a frequency of $\sim$30 kHz and with a pressure amplitude of $P_a = 1.1-1.5$ atm. Under very special conditions on experimental parameters such as the gas concentration in the liquid and the pressure amplitude, the emission of short light pulses (once per driving period) from the center of the bubble is observed. These experiments stimulated us to perform a series of studies on bubble stability.[13–18] Three types of instability mechanisms seem to be important: spherical instability, diffusive instability, and chemical instability. All of these studies are based on a Rayleigh–Plesset-like (RP) equation which provides an accurate description of the bubble wall dynamics even for strongly nonlinear oscillations. Excellent agreement with the experiments was obtained, encouraging us to rely on the RP equation also for microbubbles driven at 3 MHz. We will give an overview on RP dynamics in Sec. I. Section II shows results of calculated

sound intensities emitted into the whole spectrum as well as at the fundamental and second harmonic frequencies. Of the instability mechanisms mentioned above, only shape instabilities are important here. They will be treated in Sec. III and reveal important restrictions on useful values of driving pressure amplitudes and bubble radii. Section IV presents conclusions.

I. RAYLEIGH–PLESSET BUBBLE DYNAMICS

We will treat the dynamics and sound emission of a single bubble here; we assume that it is driven by a spatially homogeneous, standing wave field

$$P(t)=P_a \cos \omega_d t \tag{1}$$

with a frequency $\omega_d/2\pi=3$ MHz, corresponding to a period of $T=0.33$ μs, and a sound amplitude P_a between 1 and 10 atm (roughly 10^5-10^6 Pa), reflecting typical peak pressures of devices in ultrasound diagnostics. Much higher pressure amplitudes, as applied in lithotripters (cf., e.g., Ref. 19), could damage the tissue.

Let us briefly discuss the approximations we made here. In water, the chosen frequency corresponds to a sound wave length of $\lambda=2\pi c_l/\omega_d \approx 500$ μm, where $c_l=1481$ m/s is the sound velocity in water. The typical ambient radius R_0 of the microbubbles is in the range of 1–5 μm. Therefore, the approximation of spatial homogeneity is justified. In diagnostic ultrasound devices, the driving sound is not a standing wave, but a short traveling wave pulse (which is not strictly monochromatic). The corresponding spatial fluctuations of the driving pressure gradient at the location of the bubble will exert translational forces on the bubble. The resulting translational movements of the bubble are neglected (we will come back to this assumption later in this section), as well as bubble–bubble interactions, the so-called secondary Bjerknes forces.[20] Finally, pressure fluctuations due to the blood pressure (order of magnitude 0.05 atm) can also safely be neglected. In many cases, ultrasound contrast enhancers do not contain pure air bubbles, but stabilized bubbles with an albumin or saccharide coating.[3] This may lead to a shift in the resonance frequency of the bubbles as elaborated by de Jong, Church, and others.[21–24]

Under these assumptions the dynamics of the bubble radius $R(t)$ may be described by the following ordinary differential equation:[20,25]

$$R\ddot{R}+\frac{3}{2}\dot{R}^2=\frac{1}{\rho_l}(p(R,t)-P(t)-P_0) + \frac{R}{\rho_l c_l}\frac{d}{dt}(p(R,t)-P(t))-4\nu\frac{\dot{R}}{R}-\frac{2\sigma}{\rho_l R}. \tag{2}$$

Typical parameters for an air bubble in water are the surface tension $\sigma=0.073$ kg/s^2, the water viscosity $\nu=10^{-6}$ m^2/s, and density $\rho_l=1000$ kg/m^3. We use these parameters for our calculations. Only the viscosity is chosen to be larger by a factor of 3 with respect to water ($\nu=3\times10^{-6}$ m^2/s), corresponding to the value for blood. We will later see that the increased viscosity is essential for the spherical stability of the bubble at higher values of P_a. The external (ambient) pressure is $P_0=1$ atm. We assume that the pressure inside the bubble is given by a van der Waals type equation of state

$$p(R(t))=\left(P_0+\frac{2\sigma}{R_0}\right)\left(\frac{R_0^3-h^3}{R^3(t)-h^3}\right)^{\kappa} \tag{3}$$

with a (collective) van der Waals hard core radius $h=R_0/8.54$ (for air),[26] i.e., h^3 is a measure for the total excluded volume of the molecules. The bubble radius under normal conditions (ambient radius) R_0 is not uniform for the bubble population in a diagnostic suspension. The size distribution can, however, be controlled experimentally and is typically centered around 1–2 μm, with a width of about 1 μm.[3] For the effective polytropic exponent we take $\kappa\approx1$, as for the oscillation frequencies under consideration, micrometer bubbles can be treated as approximately isothermal.[27] Equation (2) can be understood as a balance equation between the excitation due to forcing (1) on the one hand and dissipative and acoustic loss processes on the other hand.

In this paper, we will denote (2) as the (modified) Rayleigh–Plesset (RP) equation, adopting a common practice in recent work on sonoluminescence.[26,28,29] Besides the pioneering work of Lord Rayleigh[25] and Plesset,[30] other researchers have contributed to Eq. (2) and a number of variations of this equation, e.g., Keller and Miksis,[31] Flynn,[32] or Gilmore.[33] Some of these variations are much more elaborate than (2). However, a detailed comparison of the solutions obtained from these equations[18,34] shows that significant deviations only occur for bubbles driven at very large pressure amplitudes ($\gtrsim$5 atm) *and* having large radii, i.e., R_0 would have to be substantially larger than for the bubbles of the present study to necessitate the use of a more complicated dynamical equation.

Let us first consider the resonance structure of the RP oscillator in the small forcing limit. To calculate the main resonance frequency we first note that for small forcing $P_a<P_0$ the contribution of the sound coupling to the bubble dynamics [the term $\propto 1/c_l$ on the RHS of Eq. (2)] is not important. In addition, if R stays large enough to ensure $R^3 \gg h^3$ (which is the case for weak driving), we can replace the van der Waals formula by an ideal gas expression. Writing $R(t)=R_0(1+x(t))$ we obtain

$$\ddot{x}=\frac{1}{\rho_l R_0^2}\left[\left(P_0+\frac{2\sigma}{R_0}\right)(1+x)^{-3\kappa-1}-\frac{P_0+P_a\cos\omega_d t}{1+x} -\frac{2\sigma}{R_0(1+x)^2}\right]-\frac{4\nu\dot{x}}{R_0^2(1+x)}-\frac{3}{2}\frac{\dot{x}^2}{1+x}. \tag{4}$$

We interpret the bracketed term on the RHS of Eq. (4) as an effective, time-dependent force $-\partial_x V(x,t)$. Integration gives the time-dependent potential

$$V(x,t)=\frac{1}{\rho_l R_0^2}\left[\frac{1}{3\kappa}\left(P_0+\frac{2\sigma}{R_0}\right)(1+x)^{-3\kappa} +(P_0+P_a\cos\omega_d t)\ln(1+x)-\frac{2\sigma}{R_0}\frac{1}{1+x}\right], \tag{5}$$

which displays a strong asymmetry in x; see Fig. 1. If $P_a=0$, the equilibrium point is $x=0$. For general P_a the minimum of the potential oscillates around this value. If $P_a>P_0$ the potential is *repulsive* for a certain fraction of the

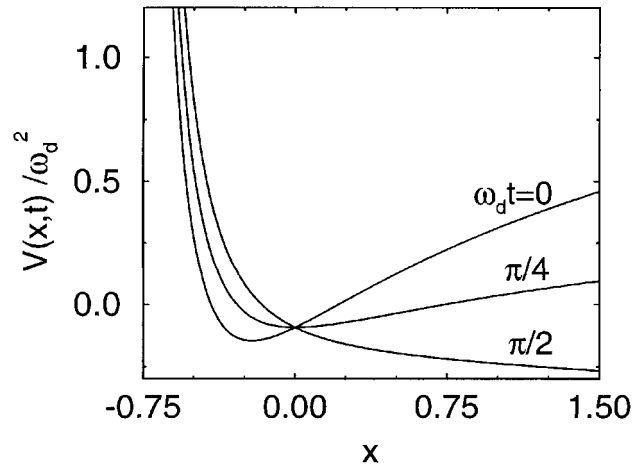


FIG. 1. Potential according to Eq. (6), nondimensionalized through dividing by ω_d^2, using $P_a=2$ atm and $R_0=1.2$ μm for three different phases $\omega_d t=0$, $\pi/4$, $\pi/2$, respectively.

period. For small P_a and thus small x we can linearize around $x=0$ and obtain a driven harmonic oscillator

$$\ddot{x}+2\gamma\dot{x}+\omega_0^2 x=\frac{P_a}{\rho_l R_0^2}\cos\,\omega_d t \tag{6}$$

with the eigenfrequency

$$\omega_0=\sqrt{\frac{1}{\rho_l R_0^2}\left(3\kappa P_0+(3\kappa-1)\frac{2\sigma}{R_0}\right)} \tag{7}$$

and the damping constant $\gamma=2\nu/R_0^2$. For fixed driving frequency $\omega_d=2\pi\cdot 3$ MHz and $\kappa=1$ the bubble oscillator is in (main) resonance if $\omega_d=\omega_0$. According to Eq. (7), this corresponds to a resonance radius of $R_0=1.23$ μm. Taking viscous damping into account, the oscillation amplitude has its maximum value at a frequency[35]

$$\omega_{\rm res}^{(1)}=\sqrt{\omega_0^2-2\gamma^2}$$
$$=\sqrt{\frac{1}{\rho_l R_0^2}\left(3\kappa P_0+(3\kappa-1)\frac{2\sigma}{R_0}\right)-\frac{8\nu^2}{R_0^4}}, \tag{8}$$

which (with $\nu=3\times10^{-6}$ m^2/s) shifts the main resonance radius to $R_0^{(1)}=1.18$ μm. The corresponding radii for the subharmonics $\omega_{\rm res}^{(1/2)}=\omega_d/2$, $\omega_{\rm res}^{(1/3)}=\omega_d/3$, and $\omega_{\rm res}^{(1/4)}=\omega_d/4$ are $R_0^{(1/2)}=2.18$ μm, $R_0^{(1/3)}=3.14$ μm, and $R_0^{(1/4)}=4.09$ μm, respectively. The harmonic $\omega_{\rm res}^{(2)}=2\omega_d$ is at $R_0^{(2)}=0.64$ μm. As we will see below, the harmonic and subharmonic resonances will also strongly affect the intensity of the emitted sound.

Figure 2 shows the time series $R(t)$ for four typical sets of parameters, as computed from (2) and (3) using a double precision, fourth-order, variable stepsize Runge–Kutta algorithm.[36] For small P_a it is trivial that the radius oscillates sinusoidally, but this can also happen for driving pressure amplitudes as large as $P_a=9$ atm, if the radius is much larger than the resonance radius, as seen in Fig. 2(a). For other parameter combinations, the bubble changes its behavior and exhibits *collapses,* characterized by short duration minima of the radius, accompanied by large accelerations [large curvature of $R(t)$]. In Fig. 2(b) we observe one (weak)

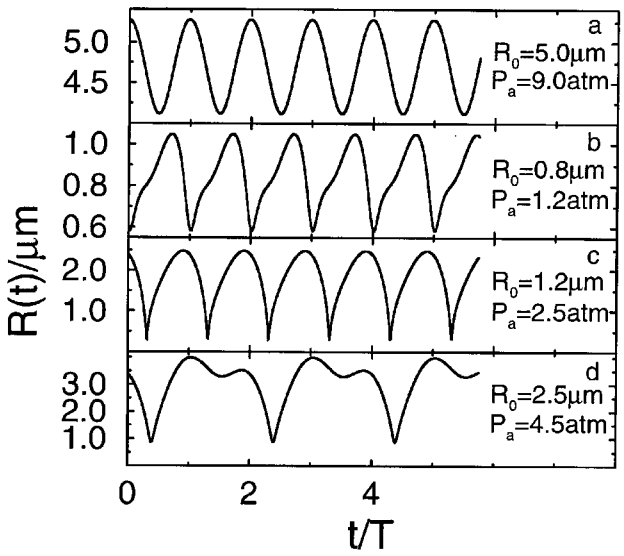


FIG. 2. The bubble radius R vs time for four different parameter pairs (R_0,P_a); from upper to lower: (a) (5 μm, 9.0 atm); (b) (0.8 μm, 1.2 atm); (c) (1.2 μm, 2.5 atm); (d) (2.5 μm, 4.5 atm).

bubble collapse per cycle which becomes stronger for larger P_a; see Fig. 2(c). For strong collapses the typical return time of the collapsing bubble is in the nanosecond range. Like the time series of most nonlinear oscillators, the bubble dynamics can period double so that a collapse only repeats every two cycles, as shown in Fig. 2(d) for a strong collapse. In other parameter regions, aperiodic behavior (''chaos'') can be observed (cf. Ref. 37). It can also be seen from Fig. 2 that our notion of strong collapse coincides well with Flynn's[32,38] definition of ''transient cavities,'' for which the ratio of maximum expansion radius and ambient radius (expansion ratio) must fulfill $R_{\rm max}/R_0\gtrsim 2$: The examples of Fig. 2(c) and (d) show rapid collapses and an expansion ratio of $\sim$2. On the other hand, Fig. 2(a) and (b) exemplifies weakly oscillating bubbles with expansion ratios near one.

For the purposes of this paper, it is instructive to com-

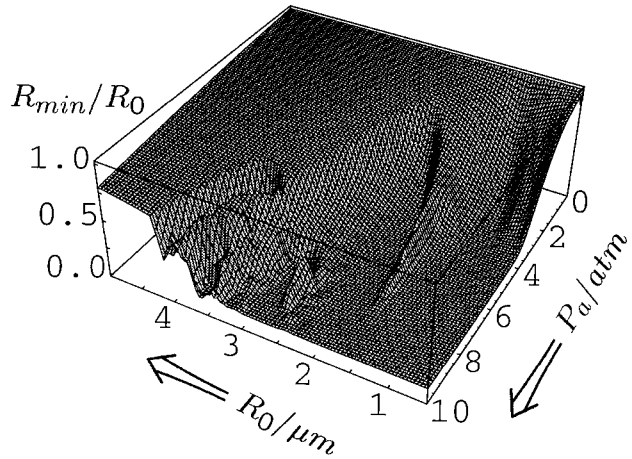


FIG. 3. Minimum radius $R_{\rm min}/R_0$ as a function of R_0 and P_a. Note that the graph is enclosed between two planes: for small P_a, the quotient $R_{\rm min}/R_0$ is nearly equal to one (weak oscillations); for large P_a and small R_0 it approaches $h/R_0=1/8.54$, with the van der Waals hard core radius h. Arrows in Figs. 3, 4, 7–9 indicate axis orientation.

pute the *minimum* radius $R_{\min}=\min_t(R(t))$ which the bubble achieves during its oscillation. This quantity is shown in Fig. 3. For this figure, as for the other 3-D plots, the displayed function was evaluated at 100×100 equidistant grid points in the P_a-R_0 plane. For weak forcing the minimum radius essentially equals the ambient radius (limit of small oscillations). However, if the driving pressure amplitude is large enough, the bubble collapse is only halted in the immediate vicinity of the smallest possible radius, i.e., the van der Waals hard core radius. The transition toward hitting the hard core radius $h=R_0/8.54$ (upon increasing P_a) is rather abrupt, forming a well-defined threshold in the P_a-R_0 plane. Obviously, this transition occurs for smaller P_a if the bubble radius is near one of the above-mentioned resonance radii. The resonances at $R_0^{(1)}=1.18$ μm and $R_0^{(1/2)}=2.18$ μm are clearly recognized in Fig. 3. As P_a becomes larger, the nonlinearities in (2) lead to broadening and a slight shift of the resonances toward smaller R_0, in accordance with earlier work on bubble dynamics.[27,31]

We now come back to our above assumptions on the pressure field. The results of this work were obtained assuming a driving by standing plane waves. Today's ultrasound diagnostics devices usually emit a bundle of traveling waves that interfere constructively to build up large pressure peaks (5–10 atm) and to achieve sufficient spatial resolution. The resulting spatial pressure gradients lead to translational forces acting on the bubble, the so called primary Bjerknes forces:[20]

$$\mathbf{F}_B(\mathbf{r},t)=V_b(t)\nabla P(\mathbf{r},t), \tag{9}$$

where $V_b(t)$ is the time-dependent volume of the bubble, and $P(\mathbf{r},t)$ is the external pressure exerted on the bubble. However, numerical computation of (9) shows that the accelerations resulting from a pressure gradient $|\nabla P|\sim P_a/\lambda$ are rather weak and that the translational velocities of the bubbles are small compared to their radial oscillation velocities.

II. SOUND EMISSION OF OSCILLATORY BUBBLES

Our focus of interest is on the sound emitted by the oscillating bubble. The far-field sound pressure at a distance $r\gg R$ from the center of the bubble can be calculated as[20,39]

$$P_s(r,t)=\frac{\rho_l}{4\pi r}\frac{d^2V_b}{dt^2}=\rho_l\frac{R}{r}(2\dot{R}^2+R\ddot{R}). \tag{10}$$

An ultrasound diagnostics device will display a picture of sound intensity, which is based on the modulus (or, equivalently, the square) of the sound pressure. Obviously, the total detected intensity will not consist exclusively of signals due to Eq. (10), but there will also be intensity components from reflections of the incoming signal (1) in the tissue. As these latter contributions depend on many peculiarities of the experimental or diagnostic setup, we do not try to model them here, but focus on the active sound radiation of the bubble.

Figure 4(a) shows the *total* sound intensity I as a function of the forcing pressure amplitude P_a and the ambient radius R_0. According to Parseval's theorem it can be calcu-

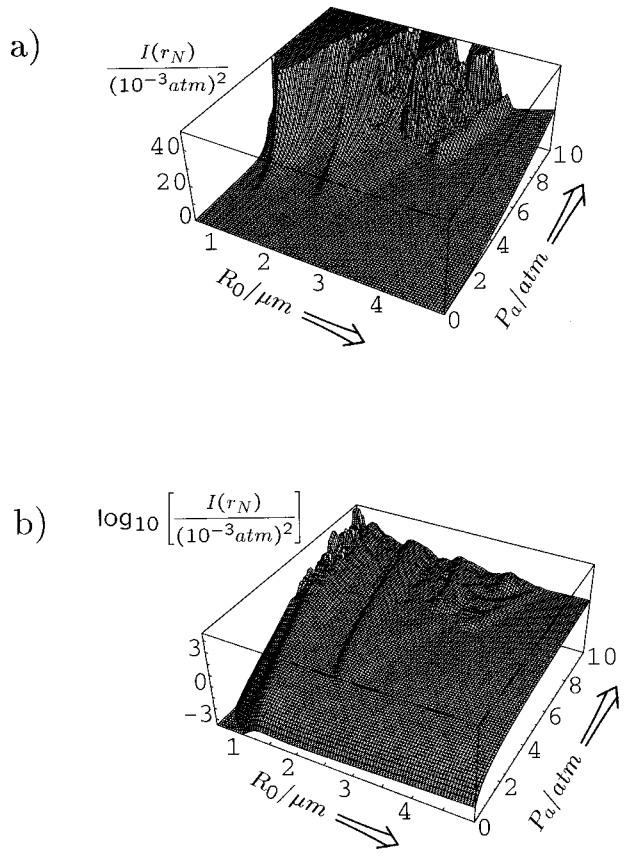


FIG. 4. (a) Total sound intensity I at a distance of $r_N=1$ cm from the bubble center as a function of R_0 and P_a. The strong correlation between this figure and Fig. 3 is obvious. Note that the graph is truncated (higher I values are not displayed) for better illustration of the resonance tongue structure. (b) shows a logarithmic plot of the same quantity. The small undulations at very large P_a on top of the resonance structure are due to numerical aliasing in the Fourier analysis and can be reduced with increasing computer power.

lated either from the sound-pressure time series $P_s(r,t)$ or from its Fourier transform $P_s(r,\omega)=\int_0^\tau P_s(r,t)\exp(i\omega t)dt$, $\tau\gg T$, namely,

$$I(r)=\frac{1}{\tau}\int_0^\tau |P_s(r,t)|^2\,dt=\frac{1}{\tau}\cdot\frac{1}{2\pi}\int_{-\infty}^{+\infty}|P_s(r,\omega)|^2\,d\omega. \tag{11}$$

We divide by the length τ of the time series ($\tau\geqslant 8T$ for all computations) and measure I in units of atm^2. The intensity is evaluated at a distance of $r_N=1$ cm from the bubble's center using standard double precision Fourier transform algorithms.[36] As I spans several orders of magnitude in our parameter regime, we also present its logarithm in Fig. 4(b).

It can be seen from Fig. 4 that for small P_a or large R_0 the bubble hardly emits any sound. Sound losses set in at a sharp threshold which is very similar to the threshold seen in Fig. 3 for the minimum radius. Moreover, comparing Figs. 3 and 4 one realizes that strong sound emission and collapsing to the hard core radius are strongly correlated (note the opposite orientations of these two graphs). This behavior is expected from Eq. (10), as a stronger collapse means larger bubble wall acceleration $\ddot{R}$ at the moment of collapse. The

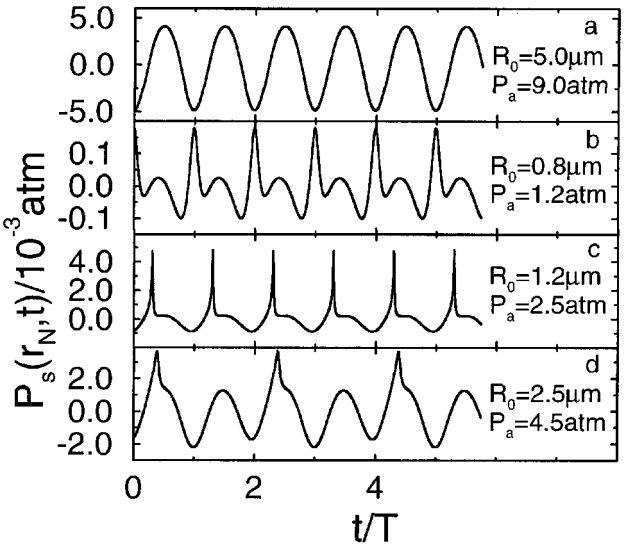


FIG. 5. Time series of the sound pressure $P_s(r_N,t)$ from (10) for the same four pairs of parameters as in Fig. 2.

resonance structure in R_0 is also clearly reflected in the emitted sound intensity.

Time series of $P_s(r_N,t)$ and their power spectra for the same four parameter pairs (R_0,P_a) as in Fig. 2 are shown in Fig. 5 and Fig. 6, respectively. If the collapse is too violent, the sound intensity is distributed in a broadband spectrum with a large fraction of intensity emitted at higher harmonics, which will be absorbed by the tissue. Therefore it is appropriate to focus on the signal at the second harmonic frequency in order to get high intensities away from the driving frequency. The intensity $I_2(r)=|P(r,\omega=2\omega_d)|^2/\tau^2$ in the second harmonic is displayed in Fig. 7. For comparison, we show the same plot also for the intensity $I_1(r)=|P(r,\omega=\omega_d)|^2/\tau^2$ of the fundamental (driving frequency) in Fig. 8.

Not surprisingly, the regions of greatest intensity of the fundamental as well as of the second harmonic coincide with

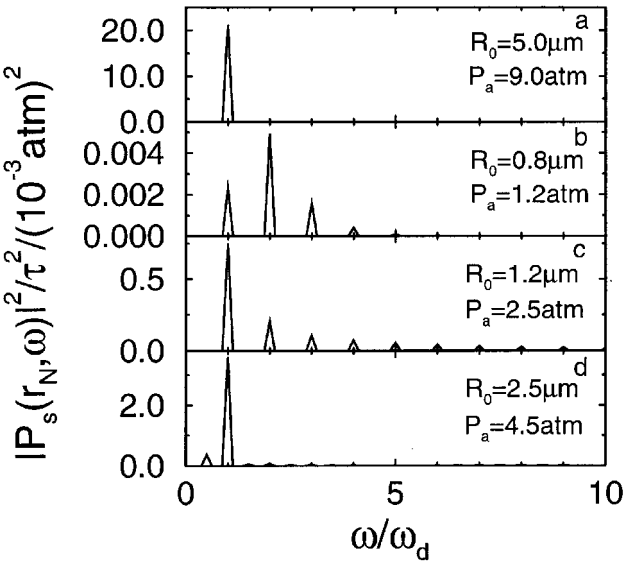


FIG. 6. Power spectra $|P_s(r_N,\omega_d)|^2$ for the same four pairs of parameters as in Fig. 2.

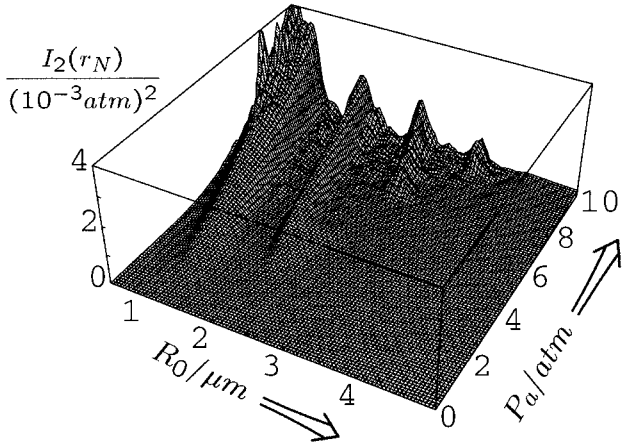


FIG. 7. Absolute intensity of the second harmonic of the sound emission (10).

those of the total intensity, i.e., they can be found at the resonance radii. The largest region and the highest maxima of second harmonic intensity occur around $R_0^{(1)}$. For small P_a the intensity is of course nearly exclusively in the driving frequency itself, which is the frequency of the small oscillations of the bubbles. Upon increasing P_a, sound is also emitted in the second harmonic mode, for some parameter combinations up to 40% of the total intensity. At even larger P_a, higher and higher modes are excited, leaving a smaller and smaller fraction of total intensity for the second mode [cf. Fig. 6(c) and (d)]. Even for large P_a, the driving frequency ω_d remains the largest component of total emitted power in spite of the strong collapse, which displays much larger peak pressures, but only lasts for extremely short periods of time.

The key question now is: How should one choose the parameters for optimal detection of the second harmonic? The answer cannot be given exclusively from the absolute intensity of the second harmonic, Fig. 7. Clearly, the signal has to have a certain absolute intensity to overcome the noise level, and therefore Fig. 7 gives important information. One important application of the second harmonic method in ultrasound diagnostics, however, relies on the contrast between intensities at fundamental and second harmonic frequencies:

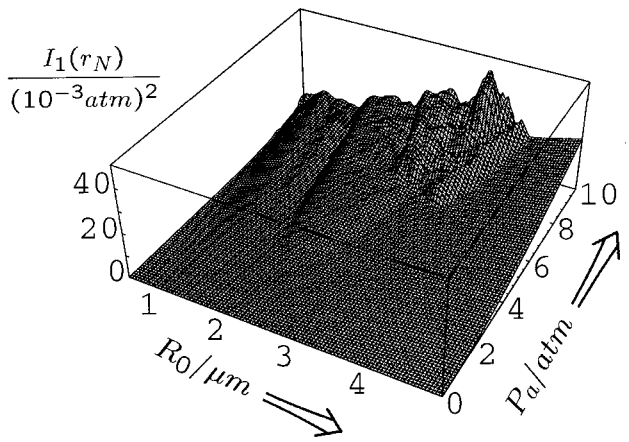


FIG. 8. Absolute intensity of the fundamental of the sound emission (10).

When injecting a bubble suspension into the vascular system, a second harmonic signal is only expected from the contrast agent. Thus detecting at 6 MHz in our example will give a bright image of the blood vessels. In a diagnostic situation, it should be possible to *switch* between this image and the scattering signal from surrounding tissue, which reflects the 3-MHz driving. But if the bubble emission signal at 3 MHz is more intense than these reflections, the vascular system will be the dominant feature in the fundamental frequency image, too. Therefore, meeting the demands of this application means to identify parameter regions where the *ratio* of second harmonic intensity to fundamental intensity I_2/I_1 is as high as possible. Figure 9 displays this quantity. It shows a distinct maximum at small $R_0 \approx 0.7-0.8$ μm close to the harmonic bubble resonance radius $R_0^{(2)}$. In this parameter region, the bubble essentially oscillates with 6 MHz instead of 3 MHz; an example for this behavior can be seen in Figs. 5(b) and 6(b). For even smaller radii, a rise in I_2/I_1 indicates other resonances. All bubbles with high I_2/I_1 ratios emit little absolute acoustic intensity (cf. Figs. 4–8).

We have presented a variety of intensity diagrams in Figs. 4–9 in order to meet the different demands of different experimental setups or diagnostic applications. Accordingly, the optimal parameter ranges for second harmonic sonography depend on the focus of interest: If the important quantity is relative intensity I_2/I_1, one would pick the maximum of Fig. 9, i.e., bubbles with $R_0 \approx 0.5-1.2$ μm and driving pressure amplitudes $P_a \approx 1-3$ atm. If absolute intensity is the key variable, one would choose the region around the main resonance radius in R_0, i.e., $R_0 \approx 1.0-1.5$ μm. Moreover, Fig. 7 suggests choosing the pressure amplitude P_a as large as possible. However, as we will show in the next paragraph, bubble shape instabilities set an upper limit on practically useful P_a. Note that a change in the driving frequency would shift the resonance radii and, consequently, also the location of regions of maximum sound emission. If, for example, ω_d is smaller, the resonances are shifted toward larger R_0; see Eq. (7).

Also, the total amount of bubbles should be large enough to guarantee a strong signal, but low enough to prevent considerable bubble–bubble interaction. Therefore, it is of primary importance to assure that a high percentage of the generated bubbles is in the correct R_0 regime. All other bubbles are essentially useless regarding the yield in the second harmonic and may even obscure the measurements.

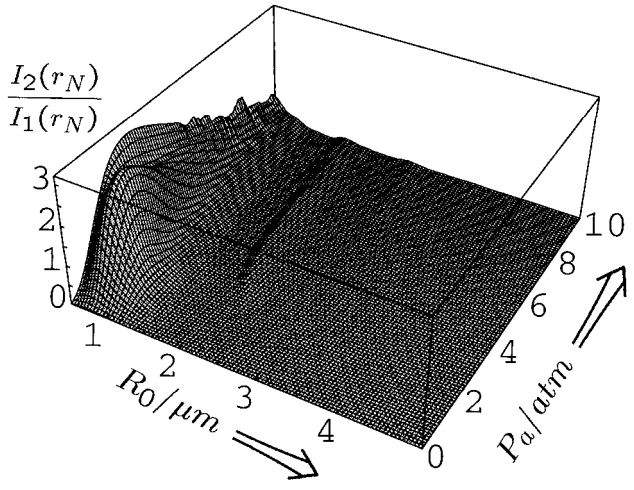


FIG. 9. Ratio I_2/I_1 of intensities of sound emission at the second harmonic and fundamental frequencies.

III. SPHERICAL STABILITY

To take advantage of the above suggested parameter regimes derived from RP dynamics, the bubbles in these domains should be *spherically stable*, i.e., stable against the growth of nonspherical bubble deformations, which could eventually lead to bubble fragmentation and a breakdown of sound emission. The corresponding stability analysis has been performed in detail in Ref. 15. We give a brief summary here. Consider a small distortion of the spherical interface $R(t)$,

$$R(t) + a_n(t) Y_n(\theta, \phi),$$

where Y_n is a spherical (surface) harmonic of degree n. An approximate linearized dynamical equation of the distortion $a_n(t)$ for each mode has been derived in Ref. 15, following the pioneering work of Prosperetti.[40] It reads (cf. Ref. 13)

$$\ddot{a}_n + B_n(t)\dot{a}_n - A_n(t)a_n = 0, \tag{12}$$

with

$$A_n(t) = (n-1)\frac{\ddot{R}}{R} - \frac{\beta_n \sigma}{\rho_w R^3} - \frac{2\nu\dot{R}}{R^3}\Big[(n-1)(n+2) + 2n(n+2)(n-1)\frac{\delta}{R}\Big], \tag{13}$$

$$B_n(t) = \frac{3\dot{R}}{R} + \frac{2\nu}{R^2}\Big[(n+2)(2n+1) - 2n(n+2)^2\frac{\delta}{R}\Big]. \tag{14}$$

Here, $\beta_n = (n-1)(n+1)(n+2)$ and δ is a viscous boundary layer cutoff,[15]

$$\delta = \min\left(\sqrt{\frac{\nu}{\omega_d}}, \frac{R}{2n}\right). \tag{15}$$

If the coefficients $A_n(t)$ and $B_n(t)$ are periodic with period T, (12) is an equation of Hill's type and instability occurs whenever the magnitude of the maximal eigenvalue of the Floquet transition matrix $F_n(T)$ of Eq. (12) is larger than one. The Floquet transition matrix $F_n(T)$ is defined by

$$\begin{pmatrix} a_n(T) \\ \dot{a}_n(T) \end{pmatrix} = F_n(T) \begin{pmatrix} a_n(0) \\ \dot{a}_n(0) \end{pmatrix}. \tag{16}$$

Now for some parameter regimes the radius is not periodic with period T and thus the coefficients $A_n(t)$ and $B_n(t)$ are not either. Therefore, rather than calculating the Floquet matrix $F_n(T)$ we calculate a transition matrix $F_n(NT)$ with a large integer N (here $N=20$) to average the dynamics. We numerically compute the eigenvalue of $F_n(NT)$. The logarithm of the maximum eigenvalue can be understood as an approximate Lyapunov exponent. If it is positive, the mode $a_n(t)$ grows exponentially and the bubble is unstable toward the corresponding mode of shape oscillation. In Fig. 10(b) and (c) we show the resulting stability diagrams for the sec-

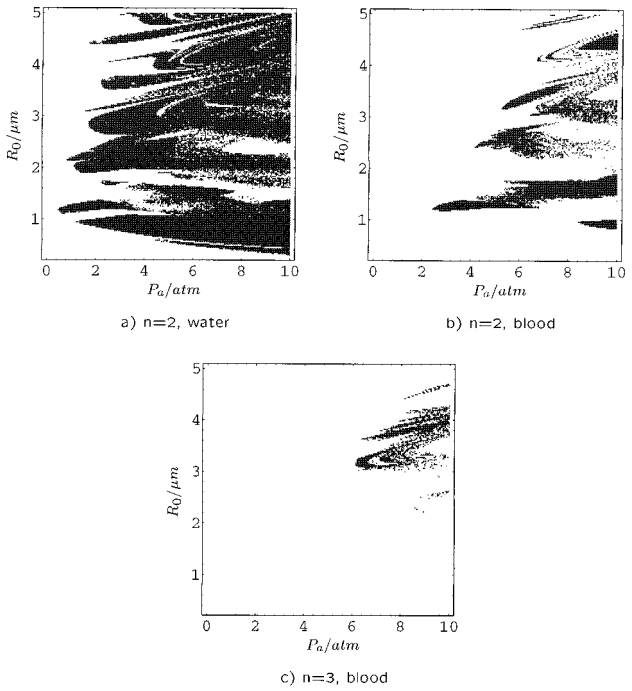


FIG. 10. Stability diagram for the $n=2$ mode for bubbles in water (a) and in blood (b) and for the $n=3$ mode for bubbles in blood (c) according to the Floquet multipliers computed from Eq. (16). In the white regions the bubbles are parametrically stable toward perturbations of the corresponding mode, in the dark regions they are parametrically unstable. Bubbles in water are much less stable than those in blood and higher modes a_n with $n \geqslant 3$ tend to be more stable than the second mode a_2. We stress that details of these stability diagrams depend on our approximation (12)–(15) as well as on the choice of the parameters of the liquid and the averaging time which is $20T$ here. Thus they should only be considered as a reflection of the general trend.

ond and the third mode ($n=2$ and $n=3$, respectively). Bubbles are shape unstable in the dark regions of the P_a-R_0 plane, and shape stable in the white areas. Generally, the $n=2$ mode is the most unstable one, but there are regimes where this does not hold.

The most pronounced features of these stability diagrams are ''tongues'' of instability. In the low P_a regime Eq. (12) reduces to a Mathieu equation; in this case the tongues of instability are the well-known Mathieu tongues, as demonstrated in Ref. 13. Large viscosity strongly damps out this tongue structure, as seen from comparing the stability diagrams for water and blood (different viscosities) in Fig. 10(a) and (b). In some regions of parameter space, stable and unstable points seem to be mixed erratically. This is due to long periodic or chaotic bubble dynamics for these P_a-R_0 combinations, for which the results of our stability analysis over $20T$ depend sensitively on the initial conditions, so that for slightly deviating parameters the stability behavior may be completely different.

As we learn from this figure, bubbles in the R_0 regime of maximum (absolute) yield in the second harmonic become spherically unstable for drivings with $P_a \gtrsim 2.5$ atm. Below, the bubbles are stable in blood (due to its enhanced viscosity), whereas for water these bubbles are unstable with respect to the a_2 mode even at $P_a \approx 0.5$ atm. Note that it is risky even to be close to an unstable regime, as the bubble may diffusionally shrink or grow into these regimes. This suggests that experiments with bubble suspensions in water will probably give misleading results (with respect to clinical applications) and should rather be carried out in blood or a fluid of correspondingly enhanced viscosity. In our calculations, we have assumed a viscosity of $\nu = 3 \times 10^{-6}$ m^2/s, which is at the low end of typical measured blood viscosities ($\sim 3-5 \times 10^{-6}$ m^2/s; Ref. 41). Thus we have determined lower bounds of the instability thresholds in Fig. 10 and the actual regions of stability may be somewhat larger. Note also that stability will be enhanced when lower driving frequencies are used as this leads to a larger effective damping of surface oscillations (cf. Ref. 13).

Stability analysis shows that very large driving amplitudes (say 10 atm) will not help provide large response signals from the suspended bubbles. Instead, it may be useful to limit the amplitudes in the bubble regions to $\lesssim 2.5$ atm.

The predictions for the region of optimal I_2/I_1 intensity ratio remain virtually unaltered, as there is very little overlap of this region with the instability tongues. Indeed, according to the regime of large I_2/I_1 identified above, shape instability in this regime is to be expected only in a tiny area of parameter space at $R_0 \approx 1.2$ μm and $P_a \approx 3$ atm [see Fig. 10(b)].

Besides spherical instability, diffusive instability and chemical instability are also matters of concern, as pointed out in detail for 26-kHz forced bubbles in Refs. 15 and 16. Both types of instabilities will only be important on long timescales (milliseconds or longer) where our approximation of a stable standing wave is not appropriate anyhow. This is why we postpone the discussion of these instabilities to future work.

IV. SUMMARY AND CONCLUSIONS

Bubble suspensions as contrast enhancers in ultrasound diagnostics are now state of the art. Improving the image quality by detecting the second harmonic of the driving frequency in the emitted sound spectrum is likely to lead to their further acceptance, as the advantages compared to conventional methods become even more pronounced. We have identified regions in parameter space, i.e., values for the driving pressure amplitude P_a and the ambient bubble radius R_0, where relatively high sound intensities at the second harmonic frequency are to be expected. These regions are intimately connected to the resonance structure of the bubble oscillator and to the collapse dynamics of the bubble. The best suited parameter regime to achieve a high absolute second harmonic intensity I_2 is located around the main bubble resonance radius, i.e., $R_0 = 1.0-1.5$ μm if a fixed driving frequency of 3 MHz is used. Requiring bubble stability toward nonspherical perturbations limits useful driving pressures to a maximum of about 2.5 atm, if the bubbles oscillate in blood [cf. Fig. 10(b)]. Here, we employed the previously specified values for ρ_l, c_l, ν, σ to model the average properties of blood. Bubbles in this parameter range are stable in fluids with (at least) three times higher viscosity than water.

The intensity ratio I_2/I_1 is important for diagnostic purposes (switching between ''background'' and contrast agent

images). It is optimized for bubbles around $R_0 \approx 0.8$ μm and $P_a \approx 1-3$ atm. Note that, again, there is an upper limit to the strength of optimal driving.

We therefore suggest not to use very high ($\geqslant 2.5$ atm) pressure amplitudes; a gentler driving may lead to a better image quality. Also, the bubble radii should be somewhat smaller than those dominant in bubble suspensions used today (e.g., SHU 508 A with a radius distribution peak at ≈ 1.4 μm, Ref. 3), if maximum efficiency at 3-MHz driving frequency is to be achieved. A narrower distribution around the peak (i.e., more uniform radii) would, of course, further amplify the sound signal.

ACKNOWLEDGMENTS

The authors thank T. Matula for valuable comments on the manuscript. This work has been supported by the DFG through its SFB185.

[1] R. Gramiak and P. M. Shah, Invest. Radiol. **3**, 356 (1968).
[2] R. Y. Nishi, Acustica **33**, 65 (1975).
[3] See the articles in *Advances in Echo Imaging Using Contrast Enhancement*, edited by N. C. Nanda and R. Schlief (Kluwer Academic, Dordrecht, 1993).
[4] C. M. Sehgal and J. F. Greenleaf, J. Acoust. Soc. Am. **72**, 1711 (1982).
[5] H. A. H. Jongen, J. M. Thijssen, M. van den Aarssen, and W. A. Verhoef, J. Acoust. Soc. Am. **79**, 535 (1986).
[6] B. A. Schrope and V. L. Newhouse, Ultrasound Med. Biol. **19**, 567 (1993).
[7] P. N. Burns, Clin. Radiol. Suppl. **51**, 50 (1996).
[8] B. Ward, A. C. Baker, and V. F. Humphrey, J. Acoust. Soc. Am. **101**, 143 (1997).
[9] D. F. Gaitan, L. A. Crum, R. A. Roy, and C. C. Church, J. Acoust. Soc. Am. **91**, 3166 (1992).
[10] B. P. Barber and S. J. Putterman, Nature (London) **352**, 318 (1991).
[11] R. Hiller, K. Weninger, S. J. Putterman, and B. P. Barber, Science **266**, 248 (1994).
[12] For reviews, see L. A. Crum, Phys. Today **47**, 22 (1994); S. J. Putterman, Sci. Am. **272**, 32 (1995); D. Lohse, Phys. Blätt, **51**, 1087 (1995).
[13] M. P. Brenner, D. Lohse, and T. F. Dupont, Phys. Rev. Lett. **75**, 954 (1995).
[14] M. P. Brenner, D. Lohse, D. Oxtoby, and T. F. Dupont, Phys. Rev. Lett. **76**, 1158 (1996).
[15] S. Hilgenfeldt, D. Lohse, and M. P. Brenner, Phys. Fluids **8**, 2808 (1996).
[16] D. Lohse, M. Brenner, T. Dupont, S. Hilgenfeldt, and B. Johnston, Phys. Rev. Lett. **78**, 1359 (1997); M. P. Brenner, S. Hilgenfeldt, and D. Lohse, ''Why air bubbles in water glow so easily,'' in *Nonlinear Physics of Complex Systems—Current Status and Future Trends*, edited by J. Parisi, S. C. Müller, and W. Zimmermann, Lecture Notes in Physics (Springer-Verlag, Berlin, 1996), p. 79.
[17] M. P. Brenner, S. Hilgenfeldt, D. Lohse, and R. Rosales, Phys. Rev. Lett. **77**, 3467 (1996).
[18] S. Hilgenfeldt, M. P. Brenner, S. Grossmann, and D. Lohse, ''Analysis of Rayleigh-Plesset dynamics for sonoluminescing bubbles,'' J. Fluid Mech. (submitted).
[19] C. C. Church, J. Acoust. Soc. Am. **86**, 215 (1989).
[20] C. E. Brennen, *Cavitation and Bubble Dynamics* (Oxford U. P., Oxford, 1995).
[21] N. deJong, F. Ten Cate, C. T. Lancee, T. C. Roelandt, and N. Bom, Ultrasonics **29**, 324 (1991).
[22] N. deJong, L. Hoff, T. Skotland, and N. Bom, Ultrasonics **30**, 95 (1992).
[23] C. C. Church, J. Acoust. Soc. Am. **97**, 1510 (1995).
[24] Z. Ye, J. Acoust. Soc. Am. **100**, 2011 (1996).
[25] Lord Rayleigh, Philos. Mag. **34**, 94 (1917); M. Plesset, J. Appl. Mech. **16**, 277 (1949); W. Lauterborn, J. Acoust. Soc. Am. **59**, 283 (1976).
[26] R. Löfstedt, B. P. Barber, and S. J. Putterman, Phys. Fluids A **5**, 2911 (1993).
[27] M. Plesset and A. Prosperetti, Annu. Rev. Fluid Mech. **9**, 145 (1977).
[28] B. P. Barber and S. J. Putterman, Phys. Rev. Lett. **69**, 3839 (1992).
[29] R. Löfstedt, K. Weninger, S. J. Putterman, and B. P. Barber, Phys. Rev. E **51**, 4400 (1995).
[30] M. Plesset, J. Appl. Mech. **16**, 277 (1949).
[31] J. B. Keller and M. J. Miksis, J. Acoust. Soc. Am. **68**, 628 (1980).
[32] H. G. Flynn, J. Acoust. Soc. Am. **58**, 1160 (1975).
[33] F. R. Gilmore, Hydrodynamics Laboratory, California Institute of Technology, Pasadena, Report **26-4** (1952).
[34] G. J. Lastman and R. A. Wentzell, J. Acoust. Soc. Am. **69**, 638 (1981); G. J. Lastman and R. A. Wentzell, J. Acoust. Soc. Am. **71**, 835 (1982).
[35] L. D. Landau and E. M. Lifshitz, *Mechanics* (Pergamon, Oxford, 1960).
[36] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in FORTRAN* (Cambridge U. P., London, 1992).
[37] W. Lauterborn and U. Parlitz, J. Acoust. Soc. Am. **84**, 1975 (1988).
[38] H. G. Flynn and C. C. Church, J. Acoust. Soc. Am. **84**, 985 (1988).
[39] L. D. Landau and E. M. Lifshitz, *Fluid Mechanics* (Pergamon, Oxford, 1987).
[40] A. Prosperetti, Q. Appl. Math. **34**, 339 (1977), see also Ref. 20 and H. W. Strube, Acustica **25**, 289 (1971); M. Plesset, J. Appl. Phys. **25**, 96 (1954); G. Birkhoff, Q. Appl. Math. **12**, 306 (1954); A. Eller and L. A. Crum, J. Acoust. Soc. Am. Suppl. 1 **47**, 762 (1970).
[41] H. Landolt and R. Börnstein, *Zahlenwerte und Funktionen aus Physik and Chemie* (Springer-Verlag, Berlin, 1969).

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.

An improved approximation for wave propagation above an impedance ground in a medium with a linear sound-speed profile

Y. L. Li
Intel Corporation, CH5-137, 5000 West Chandler Boulevard, Chandler, Arizona 85226-3699

Michael J. White[a] and Jianfeng Tai[b]
U.S. Army Construction Engineering Research Laboratory, P.O. Box 9005, Champaign, Illinois 61826-9005

(Received 13 December 1994; revised 7 April 1997; accepted 21 April 1997)

The atmospheric profile whose sound speed varies linearly with height is simple in concept, but leads to complications when solving for the sound pressure. Its effects are commonly approximated by a similar profile whose squared refractive index is a linear function of height. In this paper, the validity of the approximation has been examined for sound propagation above an impedance ground and an improved approximation is given. The new approximation is based on a WKB-type analysis using Airy functions as the general solution to the wave equation. Numerical examples show that this new approximation more adequately models the sound field for the linear sound-speed profile for a wide range of detector heights and gradients and for both upward and downward refracting atmospheres. © *1997 Acoustical Society of America.* [S0001-4966(97)02708-2]

PACS numbers: 43.28.Fp, 43.20.Fn [LCS]

INTRODUCTION

The prediction of sound propagation above an impedance ground is an important topic in the area of airborne sound propagation. The atmospheric profile whose squared refraction index varies linearly with height is a commonly used replacement for an atmosphere whose sound *speed* is a linear function of height.[1–8] The solutions using residue theory were provided by Pierce[1] and later by Berry and Daigle.[2] These solutions have been used to study and predict sound propagation within an upward refracting atmosphere.[3–6] These works are also extended to the downward refracting atmosphere.[7,8] All of the solutions rest on an assumption that a linear variation of the sound-speed profile with height can be approximated by a squared refraction index profile that is also linear in height. In this paper, some attention will be given to the approximation.

Recently, the modified Airy function (MAF) solution of the wave equation was revisited.[9] The MAF method is a WKB-type approximation using the Airy functions as the general solutions of the wave equation.[9,10] This method has been proven very powerful for computing the height-dependent Green's functions due to a point source embedded in either downward or upward refracting media.[11] Using the MAF method, the height-dependent Green's functions are presented in this paper for cases where the sound-speed profile is linear in height.

A brief overview of the conventional approximation is given in Sec. I, where the height-dependent Green's functions are given for a point source in a squared refraction index profile that is linear in height. The MAF method is used in Sec. II to obtain the height-dependent Green's functions corresponding to a point source embedded in atmospheres with a sound-speed profile linear in height. Numerical comparisons are shown in Sec. III of the conventional approximation, the improved approximation, and direct numerical integration.

I. CONVENTIONAL APPROXIMATION

The Green's function $G(\mathbf{r},\mathbf{r}_s)$ due to a time harmonic $(e^{-i\omega t})$ point source at $\mathbf{r}_s$ in a medium with a sound-speed profile that is linear in height can be cast in terms of a Sommerfeld-type integral[1,2,12]

$$G(\mathbf{r},\mathbf{r}_s)=\frac{1}{2\pi}\int_0^\infty \hat{G}(z,z_s;k)kJ_0(k\rho)dk, \tag{1}$$

where J_0 is the Bessel function of the first kind and ρ is the horizontal distance between the source and detector. The

[a]Present address: Applied Research Associates, 5941 S. Middlefield Rd., Ste. 100, Littleton, CO 80123.
[b]Present address: Material & Distribution Management, BDM International, Carnegie Hall Tower, 152 W. 57th St., New York, NY 10019.

sound-pressure field is $p(\mathbf{r},\mathbf{r}_s)=4\pi p_0 SG(\mathbf{r},\mathbf{r}_s)$, where p_0 is the pressure at distance $S=1$ m away from a point source radiating into an unbounded medium. Here, the units of G and $\hat{G}$ are, respectively, m^{-1} and m. The height-dependent Green's function $\hat{G}(z,z_s;k)$ is the solution of the Fourier-transformed Helmholtz equation:

$$\frac{d^2\hat{G}(z,z_s;k)}{dz^2}+[\omega^2/c^2(z)-k^2]\hat{G}(z,z_s;k)=-\delta(z-z_s), \tag{2}$$

the boundary condition at the ground surface

$$\left[\frac{d\hat{G}(z,z_s;k)}{dz}+i\left(\frac{\omega}{c_0}\right)\beta\hat{G}(z,z_s;k)\right]\Bigg|_{z=0}=0, \tag{3}$$

and radiation conditions at infinity, with angular frequency ω and β is the normalized admittance of the ground. The sound-speed profile of the atmosphere we seek to model is $c(z)=c_0(1+az)$, with c_0 representing the sound speed at the surface and (a) is a real parameter with unit m^{-1}. The index of refraction for this profile is $n(z)=c_0/c(z)=(1+az)^{-1}$. Equation (3) describes the influence of the ground and implies a ground surface that reacts locally to incident waves. In an unbounded medium of constant speed c_0, the solution $\hat{G}=(i/2k_z)\exp(ik_z|z-z_s|)$, $k_z=\sqrt{k_0^2-k^2}$, $k_0=\omega/c_0$ to Eq. (2) yields $G(\mathbf{r},\mathbf{r}_s)=\exp(ik_0R)/4\pi R$, with $R=\sqrt{\rho^2+(z-z_s)^2}$, as expected.

The differential equation [Eq. (2)] has general solutions composed of Bessel functions with complex order and real argument.[13–15] We found it very difficult to evaluate these Bessel functions for cases of interest. Previous attacks on this problem avoid directly solving Eq. (2) and aim instead at a similar differential equation with solutions which are more widely studied and easier to evaluate numerically. For small values of (az), the squared refraction index profile was approximated by a linear function of height,[1]

$$\omega^2/c^2(z)=k_0^2n^2(z)\approx k_0^2(1-2az). \tag{4}$$

Substituting Eq. (4) into Eq. (2), the problem equation can be rewritten as

$$\frac{d^2\hat{G}(z,z_s;k)}{dz^2}+[k_0^2(1-2az)-k^2]\hat{G}(z,z_s;k)=-\delta(z-z_s). \tag{5}$$

This differential equation has general solutions composed of Airy functions.[13] In our experience, the evaluation of the Airy functions is much simpler than evaluation of the Bessel functions of complex order and real argument. Upon substitution of the general solution into Eq. (5), application of the boundary conditions at the ground and infinity, and enforcement of continuity of the field at the source height, the height-dependent Green's function for upward refracting atmospheres $(a<0)$, is

$$\hat{G}(z,z_s;k)=2\pi e^{i\pi/6}l\ \mathrm{Ai}[(\tau-z_>/l)e^{i2\pi/3}]\{\mathrm{Ai}(\tau-z_</l)+\mathscr{R}^{-1}(k)\mathrm{Ai}[(\tau-z_</l)e^{i2\pi/3}]\}, \tag{6}$$

where $\tau=(k^2-k_0^2)l^2$, $l=(2|a|k_0^2)^{-1/3}$, $z_>=\max(z,z_s)$, $z_<=\min(z,z_s)$, and $\mathrm{Ai}'(\psi)=(d/d\psi)\mathrm{Ai}(\psi)$. For downward refracting atmospheres $(a>0)$, the height-dependent Green's function has the formulation

$$\hat{G}(z,z_s;k)=2\pi e^{i\pi/6}l\ \mathrm{Ai}(\tau+z_>/l)\{\mathrm{Ai}[(\tau+z_</l)e^{i2\pi/3}]+\mathscr{R}(k)\mathrm{Ai}(\tau+z_</l)\}. \tag{7}$$

In Eqs. (6) and (7) the coefficient $\mathscr{R}(k)$ of the ground reflected wave is defined as

$$\mathscr{R}(k)=-\frac{\mathrm{Ai}'(\tau e^{i2\pi/3})e^{i2\pi/3}\pm ik_0\beta l\ \mathrm{Ai}(\tau e^{i2\pi/3})}{\mathrm{Ai}'(\tau)\pm ik_0\beta l\ \mathrm{Ai}(\tau)}, \tag{8}$$

where β is the normalized admittance of the ground. The upper signs in Eq. (8) correspond to downward refraction, and the lower signs otherwise apply. Equations (6)–(8) complete the *exact* solution of Eq. (5) for linear variation of $n^2(z)$ with height, and the *approximate* solution for linear $c(z)$.

II. IMPROVED APPROXIMATION

Instead of pursuing the replacement of the linear sound-speed profile by the linear squared refraction index profile, the new approximation we present is based on directly solving the Fourier-transformed Helmholtz equation with the MAF solution procedure.[9,11] Following the MAF method, the general solutions of Eq. (2) can be written as

$$\hat{G}(z,z_s;k)=\{C_1\ \mathrm{Ai}[\alpha(z)]+C_2\ \mathrm{Ai}[\alpha(z)e^{i2\pi/3}]\}/\sqrt{\alpha'(z)}, \tag{9}$$

where Ai represents the Airy functions. The general solution given in Eq. (9) is valid (although approximate) everywhere, including the turning point.[9,10] In Eq. (9), $\alpha(z)$ is a scale function to be evaluated below.

Using Eq. (9) and the boundary conditions, the height-dependent Green's function for upward refracting atmospheres above impedance ground can be written

$$\hat{G}(z,z_s;k)=2\pi e^{i\pi/6}\frac{\mathrm{Ai}(\alpha_>e^{i2\pi/3})}{\sqrt{\alpha'_>\alpha'_<}}[\mathrm{Ai}(\alpha_<)+\mathscr{R}^{-1}(k)\mathrm{Ai}(\alpha_<e^{i2\pi/3})], \tag{10}$$

and following similar procedures for downward refracting atmospheres, one obtains for the Green's function

$$\hat{G}(z,z_s;k)=2\pi e^{i\pi/6}\frac{\mathrm{Ai}(\alpha_>)}{\sqrt{\alpha'_>\alpha'_<}}[\mathrm{Ai}(\alpha_<e^{i2\pi/3})+\mathscr{R}(k)\mathrm{Ai}(\alpha_<)], \tag{11}$$

where

$$\mathscr{R}(k)=-\frac{\mathrm{Ai}'(\alpha_0e^{i2\pi/3})\alpha'_0e^{i2\pi/3}+ik_0\tilde{\beta}\ \mathrm{Ai}(\alpha_0e^{i2\pi/3})}{\mathrm{Ai}'(\alpha_0)\alpha'_0+ik_0\tilde{\beta}\ \mathrm{Ai}(\alpha_0)}, \tag{12}$$

with

$$\tilde{\beta}=\beta+\frac{i\alpha_0''}{2k_0\alpha_0'}, \tag{13}$$

where primed quantities refer to derivatives with respect to their arguments, and subscripted arguments denote evaluation of the function for argument at the subscript. For example, $\alpha_0''=d^2\alpha(z)/dz^2|_{z=0}$ and $\alpha_<=\alpha(z_<)$. In Eqs. (9)–(13), the function $\alpha(z)$ is defined as

$$\alpha(z)\equiv\mp\left[\frac{3}{2}\int_{\min(z,z_t)}^{\max(z,z_t)}\sqrt{\pm\Omega^2(z)}dz\right]^{2/3}, \tag{14}$$

with the upper or lower signs chosen to make the radicand positive. In Eq. (14), Ω^2 is equal to the bracketed factor in the second term of Eq. (2), i.e.,

$$\Omega^2(z)=\omega^2/c^2(z)-k^2, \tag{15}$$

and z_t is the turning point, $\Omega(z)|_{z=z_t}=0$. For Ω a general function of height, the closed-form expression of the scale function $\alpha(z)$ is usually not obtainable. Thus numerical evaluation of the integral in Eq. (14) often becomes a necessity. Fortunately, however, for linear $c(z)$, the integrand

$$\Omega^2(z)=k_0^2/(1+az)^2-k^2 \tag{16}$$

is relatively simple. For conciseness, let $s=(k_0/k)/(1+az)$. Then, at the turning point, $s=1$ and the closed-form expression for $\alpha(z)$ is[11]

$$\alpha(z)=\begin{cases}\left\{\dfrac{3k_0}{2|a|}\left[\arcsin(s)-\dfrac{\pi}{2}+\sqrt{s^{-2}-1}\right]\right\}^{2/3}, & s\leqslant 1,\\ -\left\{\dfrac{3k_0}{2|a|}\left[\ln(s+\sqrt{s^2-1})-\sqrt{1-s^{-2}}\right]\right\}^{2/3}, & s\geqslant 1.\end{cases} \tag{17}$$

Note that computations involving the new approximation in Eqs. (10) and (11) involve very little additional effort over that of Eqs. (6) and (7), as only the arguments to the Airy functions are modified.

III. COMPARISON

In this section, some typical cases will be examined using the two approximations (the exact solution for the linear squared index profile and the MAF approximation for the linear speed profile). These cases are also computed by numerical evaluation of Eq. (1), using the transmission-line formulation to solve Eqs. (2) and (3) with a linear sound-speed profile of thin, constant-speed layers to serve as a benchmark.[12,15,16] The transformed height-dependent Green's functions in Eq. (1) are computed by both the conventional approximation Eqs. (6) and (7) and the MAF approximation Eqs. (10) and (11). In all cases, spectral integration of Eq. (1) were performed within the CFFP, so that differences in results are due solely to the differences between Green's functions. In the following examples, we display results from our computations in terms of the relative gain in sound-pressure level

$$L=20\log(|p/p_0|),$$

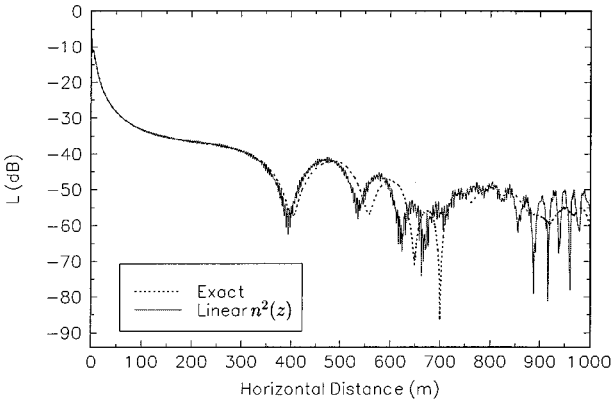


FIG. 1. Comparison of curves of relative sound-pressure level using linear $n^2(z)$ and the ''exact'' solution from the CFFP model. A point source is embedded in a downward refracting atmosphere above an impedance ground. In the atmosphere, the sound-speed profile is a linear function of height. The refraction parameter (a) is equal to 10^{-3} m^{-1}. The ground flow resistivity is 200 kN s/m^4. The frequency is 100 Hz.

where p_0 is the reference sound pressure at 1 m in the absence of the ground and without refraction.

Figure 1 shows curves of the relative sound-pressure level as a function of horizontal distance. In this example, the sound speed is given to be a linear function of height and the refraction parameter is $a=10^{-3}$ m^{-1}, such that the atmosphere is downward refracting. A point source is located at $x=y=0$, $z=2$ m, in Cartesian coordinates. The sound speed at the surface of the ground is $c_0=344$ m/s, the height of the detectors is 1 m, and the frequency is 100 Hz. The sound-speed gradient, $c_0a=0.344$ s^{-1}, is larger than what typically occurs over any great height in the real atmosphere, but can be frequently found near the ground. The value of the normalized admittance of the ground is $(0.0249-i0.0302)$, which is obtained from the Delany–Bazley model.[17] In that model, the flow resistivity was set equal to 200 kN s/m^4. The relation $az\sim 10^{-3}\ll 1$ ensures that the given linear sound-speed profile and the approximate linear squared index profile are practically identical near the ground. From Fig. 1, we see that results from Eq. (7) and the thin layer solution (denoted ''exact'' in the figures) depart at

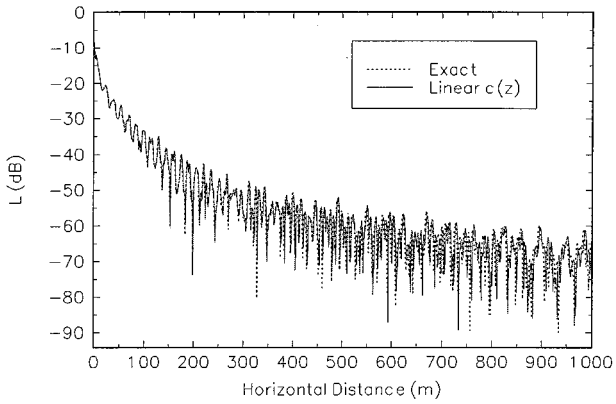


FIG. 2. Same as Fig. (1), except $a=10^{-1}$ m^{-1}. Only comparisons between the new MAF approximation and the ''exact'' solution from the CFFP model are shown.

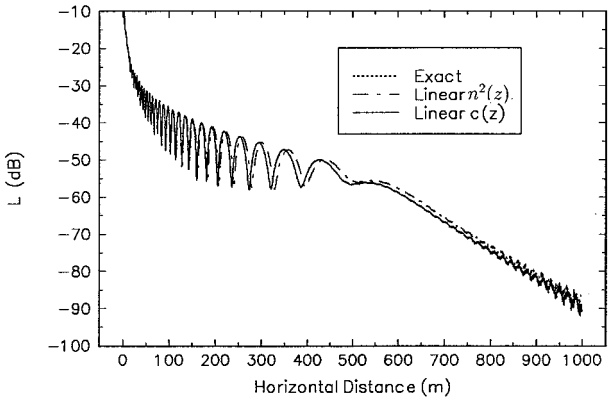


FIG. 3. Same as Fig. (1), except $a=-10^{-3}$ m^{-1}, source height is 50 m, and receiver height is 51 m. Comparisons between the new MAF approximation for linear $c(z)$ and the ''exact'' solution from the CFFP model are shown, along with the solution for linear $n^2(z)$.

longer ranges. Note that the positions of the interference extrema are shifted in range between the two curves, and at longer distances the extrema are dissimilar. We also found the MAF approximation, Eq. (11), to be indistinguishable (within the line thickness) from the thin layer solution for this case.

In Fig. 2, under conditions otherwise identical to that of Fig. 1, the sound-speed gradient is set equal to a very large value, $a=0.1$ m^{-1}. The similarity of the two curves demonstrates that the MAF approximation holds in cases where the conventional approximation [Eq. (7)] is less appropriate ($az\gtrsim 1$). The results of Eq. (7) are much less in agreement with the thin layer solution (labeled as ''exact'') but are suppressed from Fig. 2, as they obscure the comparison.

The computations for Fig. 3 were made under the same conditions as Fig. 1, but with an upward refracting atmosphere where $a=-10^{-3}$ m^{-1}, the source height set equal to 50 m, and detector height equal to 51 m. These parameters imply $|az|\sim 0.05$. Note that the linear $n^2(z)$ results differ from the other two methods by a few decibels beyond 250 m. As in the other cases, the MAF approximation gave more accurate results, in this case virtually identical to the desired thin layer results (''exact'').

When we examined the case of Fig. 1 with only a change to upward refraction where $a=-10^{-3}$ m^{-1}, the three methods gave nearly identical results.

The examples shown are for one frequency, but the comparisons were qualitatively similar to comparisons for other frequencies that we tried. In general, higher frequencies obtained greater accuracy in the MAF method when compared with the ''exact'' solution and the differences between MAF and linear $n^2(z)$ methods were enhanced. At lower frequencies, all methods became indistinguishable.

IV. CONCLUSION

An atmosphere whose squared refraction index is linear in height is commonly used to analyze the sound pressure near the boundary in a case where the sound *speed* varies linearly with height. In this paper, the approximation has been tested by computation. Numerical examples show that this profile replacement may be less valid at longer ranges and greater values of the absolute product of the sound-speed gradient and height. In one case (Fig. 1), the approximation was inaccurate by as much as 10 dB or more beyond 500 m in distance, with $az\sim 10^{-3}\ll 1$.

An improved approximation has been introduced, based on the MAF method for solution of the Helmholtz equation. There is very little extra computational overhead required for using the method, since the conventional formulation of the Green's functions uses the same set of special functions (i.e., the Airy functions). Numerical examples show that the new approximation is excellent for either weak or strongly refracting atmospheres. The MAF method was shown to successfully model source and receiver heights near to the ground (~ 1 m) and and well above the surface (~ 50 m). In these numerical examples, only one frequency is shown, although it is representative of the results for other cases. Higher frequency often obtains greater accuracy in the MAF approximation.

[1] A. D. Pierce, *Acoustics: An Introduction to Its Physical Principles and Applications* (Acoustical Society of America, New York, 1989).
[2] A. Berry and G. A. Daigle, ''Controlled experiments of the diffraction of sound by a curved surface,'' J. Acoust. Soc. Am. **83**, 2047–2058 (1988).
[3] C. G. Don and A. J. Cramond, ''Creeping wave analysis of impulse sound propagation through a shadow boundary,'' J. Acoust. Soc. Am. **80**, 302–305 (1986).
[4] R. Raspet and S. J. Franke, ''Residue series solution of impulse noise propagation into a shadow zone,'' J. Acoust. Soc. Am. **83**, 1964–1967 (1988).
[5] S. J. Franke, R. Raspet, and C. H. Liu, ''Numerical predictions of atmospheric sound pressure levels in shadow zones,'' J. Acoust. Soc. Am. **83**, 816–820 (1988).
[6] M. West, F. Walkden, and R. A. Sack, ''The acoustic shadow produced by wind speed and temperature gradients close to the ground,'' Appl. Acoust. **27**, 239–260 (1989).
[7] K. B. Rasmussen, ''Outdoor sound propagation under the influence of wind and temperature gradients,'' J. Sound Vib. **104**, 321–335 (1986).
[8] R. Raspet, G. Baird, and W. Wu, ''Normal mode solution for low-frequency sound propagation in a downward refracting atmosphere above a complex impedance plane,'' J. Acoust. Soc. Am. **91**, 1341–1352 (1992).
[9] A. K. Ghatak, R. L. Gallawa, and I. C. Goyal, ''Modified Airy Function and WKB Solutions to the Wave Equation,'' National Institute for Standards and Technology Monograph 176 (1991).
[10] L. B. Felsen and N. Marcuvitz, *Radiation and Scattering of Waves* (Prentice-Hall, Englewood Cliffs, NJ, 1973).
[11] Y. L. Li, ''Efficient computation of the sound field above ground in horizontally stratified media using a Wentzel–Kramers–Brillouin-type approximation with Airy functions,'' J. Acoust. Soc. Am. **98**, 3405–3411 (1995).
[12] Y. L. Li, M. J. White, and S. J. Franke, ''New fast field programs for anisotropic sound propagation through an atmosphere with a wind velocity profile,'' J. Acoust. Soc. Am. **95**, 718–726 (1994).
[13] M. Abramowitz and A. I. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1965).
[14] C. L. Pekeris, ''Theory of propagation of sound in a half-space of variable sound velocity under conditions of formation of a shadow zone,'' J. Acoust. Soc. Am. **18**, 295–315 (1946).
[15] Y. L. Li, C. H. Liu, and S. J. Franke, ''Adaptive evaluation of the Sommerfeld-type integral sing the Chirp-z transform,'' IEEE Trans. Antennas Propag. **39**, 1788–1791 (1991).
[16] Y. L. Li and M. J. White, ''A note on using the fast field program,'' J. Acoust. Soc. Am. **95**, 3100–3102 (1994).
[17] M. E. Delany and E. N. Bazley, ''Acoustical properties of fibrous absorbent materials,'' Appl. Acoust. **3**, 105–116 (1970).

Symmetrical oscillation modes in choked-jet edge tones and screech from rectangular nozzles[a)]

Dan Lin and Alan Powell
Department of Mechanical Engineering, University of Houston, Houston, Texas 77204-4792

(Received 5 February 1996; revised 5 April 1997; accepted 26 April 1997)

The finding of an energetic symmetrical mode in choked-jet edge tones at relatively low-pressure ratios and small nozzle-to-edge spacings stimulated a search for the symmetrical mode in the screech of choked rectangular jets (where there is no edge). In the pressure ratio range $R=2.05 \sim 2.12$ to 2.16, near-field acoustic measurements revealed weak symmetrical modes of screech for a rectangular jet with exit aspect ratio 4.8. The wavelengths of these symmetrical sound waves are about three-quarters of the adjacent asymmetrical ones. It is proposed that the occurrence of the symmetrical mode in moderately low-aspect-ratio choked-jet screech is due to the three-dimensional effect and is sustained primarily by the feedback loop from a single acoustic source at the end of the first cell adjacent to the nozzle lip. © *1997 Acoustical Society of America.*
[S0001-4966(97)04208-2]

PACS numbers: 43.28.Ra [LCS]

INTRODUCTION

A. Symmetric mode for edge tones

The edge tone of the jet–edge system is the classic subject for the investigation of the feedback phenomenon in flows. The instability of these jets has always been considered asymmetrical (sinuous), from the classical low-speed laminar jet to the high-speed turbulent jet and choked jets (e.g., Brown, 1937; Powell, 1953a; Mascarenhas, 1992). However, recent investigations show that this is not always the case. At jet pressure ratios just below the lower limit for the expected sinuous instabilities, i.e., at pressure ratios just above choking, and at small nozzle-to-edge distances, symmetrical (varicose) modes of edge tones were found for a rectangular nozzle with aspect ratio AR=4.8 (17.3 $\times$3.6 mm). The resultant sound waves as well as the jet disturbances (instability waves) are clearly symmetrical on two sides of the jet as revealed by schlieren photography, see Fig. 1 from Lin (1992). The acoustic wavelength for this symmetric mode is about two-thirds of that of the adjacent asymmetrical mode, while the intensity for the sound generated is about 9 dB lower.

B. Symmetric mode for choked jet?

Edge tones of choked jets are known to operate in the frequency range more or less overlapping that of the screech tone of edgeless jets (Lin, 1992). Moreover, as the wedge is moved further away from the nozzle, the edge tones transition to choked-jet screech alone. The question arose as to whether the symmetrical mode could occur in the screech of a edgeless choked jet from the same nozzle and, in particular, whether an unidentified screech mode found by Kozu (Kozu, 1991) was in fact symmetric. The present paper, to answer this question, describes a brief investigation concerning the existence of the symmetric mode in choked-jet screech, which hopefully will provide a more complete understanding of the feedback phenomena in choked-jet instability and screech.

I. EXPERIMENT

Using the same nozzle (AR=4.8) as in Lin (1992) for choked-jet screech experiments, acoustic measurements by Kozu (1991) showed a discontinuity in the curve of Strouhal number $S=fU_{fe}/h$ (U_{fe}=fully expanded jet velocity, h =the smaller dimension of the rectangular nozzle exit), as seen in Fig. 2. Thus we further investigated the choked-jet screech for the nozzle that Kozu used. For this further investigation, it was necessary to rely on acoustic measurements since the sound pressure and the jet flow instability for the symmetrical mode were both too weak to be made evident by the schlieren system. This was done by obtaining signals of microphones mounted symmetrically on opposite sides of

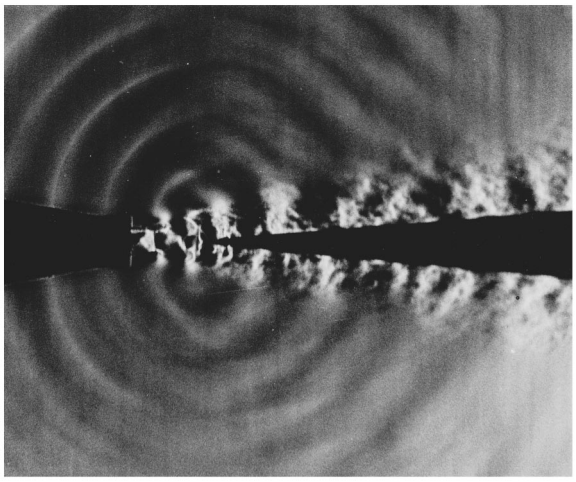


FIG. 1. Schlieren photograph for choked-jet edge tone ($R=2.36$, nozzle-to-edge spacing $x=2.8h$, wedge angle 8°).

[a)]Presented at the 128th Meeting of the Acoustical Society of America, Austin, Texas, 28 Nov.–2 Dec. 1994 [J. Acoust. Soc. Am. **96**, 3332(A) (1994)].

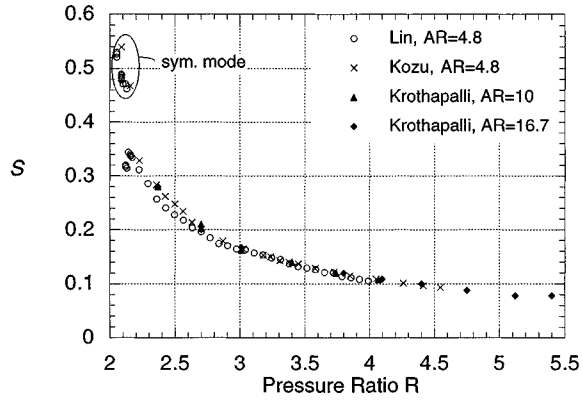


FIG. 2. Strouhal number for choked-jet screech by various investigators. A previously unidentified group of data is now determined to be of a symmetric mode.

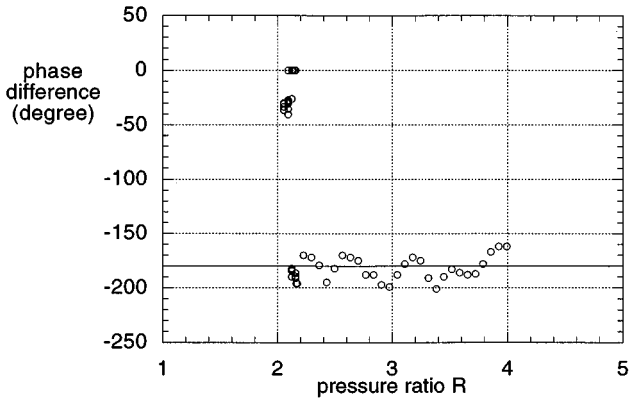


FIG. 4. Phase difference of the acoustic waves between the opposite sides of the free choked jet.

the jet. The diaphragms of the B&K 4138 microphones facing the jet were in the same plane as the nozzle exit and were 20 mm away from the outer edges of the nozzle. The signals were fed through filters for display on a dual beam oscilloscope or were fed into a CF-360 Ono Sokki fast Fourier transform (FFT) analyzer to yield values of the phase difference. At the lowest pressure ratios R [say $R<2.2$, where R = (reservoir pressure)/(ambient pressure)], the signals from the two microphones were found to be almost in phase, while at higher pressure ratios they were of the expected opposite phase as shown in Figs. 3 and 4. The sound field is evidently symmetric at the lower pressure ratios.

The acoustic wavelengths for the symmetric modes are about 25% lower than the value given by the empirical formula for asymmetrical modes $\lambda/d=K(R-R_c)^{1/2}$ with $K=5.2$ (Powell, 1953b) or 5.0 (Krothapalli, 1986). In the very narrow range between $R=2.12$ and 2.13, the mode was observed to be sometimes symmetrical, sometimes asymmetrical. The Strouhal numbers for these asymmetrical modes are noticeably lower than for the immediately adjacent ones for the regular asymmetric rectangular choked-jet screech (Fig. 2). The reason for this is unknown, and we named this region as ''transitional.''

II. DISCUSSION

Subsequently, Raman and Rice (1994) reported finding an asymmetrical mode at the fundamental frequency (f_0) and a symmetrical mode at just double the fundamental frequency ($2f_0$). The latter apparently never occurred alone. A plausible alternative explanation of Raman and Rice's finding is as follows. Let a pair of nonsinusoidal waveforms, representing the perturbation on each side of the jet, be positioned asymmetrically. Then their decomposition yields asymmetrical waveforms for f_0 and its odd harmonics ($3f_0$, $5f_0$, etc.) but symmetrical waveforms for the even harmonics ($2f_0$, $4f_0$, etc.). Hence the observed symmetrical waveform of the jet instability and the acoustic wave at $2f_0$ may well not be that of a feedback mode that could have its own independent existence.

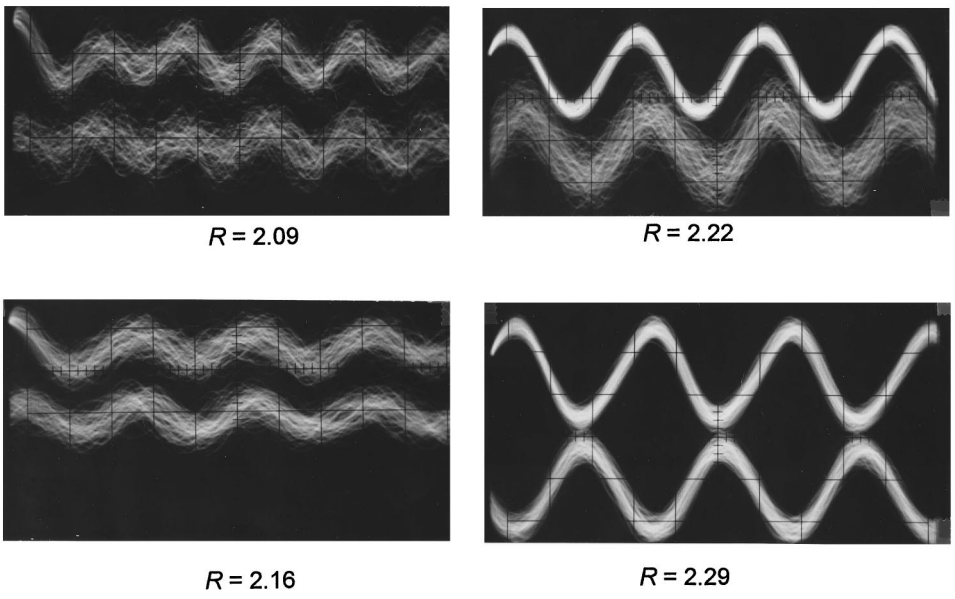


FIG. 3. Screech sound-pressure signals from opposite sides of the jet.

Rayleigh's (1945, p. 376) linear instability theory of a two-dimensional incompressible inviscid jet with a top-hat velocity profile shows that the disturbances, in terms of displacement of the jet boundary, are temporally unstable to both axisymmetric and asymmetrical disturbances. The spatial growth rates can be related to the temporal growth rate by the convection velocity of the disturbances. It is found that for long wavelengths, asymmetrical disturbances grow much faster than symmetrical ones. As disturbance wavelengths in the jet become small, the difference between the two growth rates also becomes smaller. A sophisticated two-dimensional numerical calculation for a choked jet (Umeda *et al.*, 1990) shows that the jet is stable under symmetrical disturbances, but undergoes violent sinuous motion under asymmetrical boundary conditions. This indicates that in the strictly two-dimensional case the symmetrical mode probably does not exit, or at least, is always very much weaker than its asymmetrical counterpart.

For cold supersonic round jets, the axisymmetric (varicose or toroidal) instability wave has the larger growth rate at low supersonic speeds (Tam, 1995). Indeed, experiments show that both for circular jets (e.g., Powell *et al.*, 1990) and for square jets (Umeda and Ishii, 1994), the symmetrical mode is the first to appear with increasing pressure ratio. Mattingly and Chang (1974) found that with a velocity profile of a potential core and Gaussian shear layer distributions, the circular jet is spatially unstable. For the jet near the exit, the axisymmetric mode dominates over a range of frequencies. Further downstream, however, the asymmetrical mode is the most unstable mode.

This feature may be seen in the rectangular choked-jet edge-tone case (see Fig. 1). While the jet instability wave is symmetrical between the exit and the wedge, the motion becomes asymmetric downstream of the wedge. The same phenomenon could be happening in the rectangular choked-jet screech.

Powell (1953a) suggested that the feedback mechanism of edge tones and screech are basically the same, except for the latter the sound is emitted from a phased array associated with the regularly spaced shock waves, and the choked jet offers a high acoustic impedance that separates the action on the two sides of the jet. Hence this feedback theory for screech does not require any jet disturbance asymmetry (as for low-speed edge tones), so we suspect that the mechanism for the symmetrical mode of screech is basically the same as for the asymmetrical one. When sound waves due to the acoustic sources all have the same phase at the nozzle exit, they have the maximum effect on the generation of disturbances at the nozzle exit. The sound-pressure amplitude at the nozzle exit is factored by the transmission effectiveness given by Powell (1962) as

$$\eta_t = \sum_1^n \frac{a_n}{ns} e^{-i2\pi(ns/\lambda)(1+1/M_{\text{con}})}, \tag{1}$$

where s is the cell-length spacing between the effective point sources, $s = 2\sqrt{M_j^2 - 1}$ from linear theory for lower-pressure ratios and being measured from schlieren photographs for higher-pressure ratios, in which M_j is the jet exit Mach number, λ is the acoustic wavelength, a_n is the strength for the nth source counting from the nozzle exit, $M_{\text{con}} = U_{\text{con}}/c$, with U_{con} the convection velocity of the disturbances, taken as equal to half of the fully expanded jet velocity as expected from linear theory for short disturbance wavelengths, and c is the speed of sound in the ambient medium.

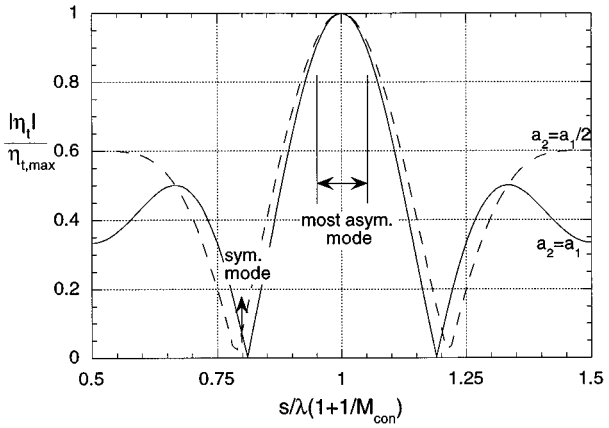


FIG. 5. The transmission effectiveness, Eq. (1) of choked-jet screech.

The value of $|\eta_t|/\eta_{t,\max}$, where $\eta_{t,\max} = \Sigma a_n/ns$, gives the selectivity of the feedback transmission (Powell, 1962), and the screech frequency is expected to occur near the peak value of $|\eta_t|/\eta_{t,\max}$. The same result would occur if the sound source is considered as a continuous phased array associated with the shock-cell structure as proposed by Tam (1995). The plausible values $a_2 = a_1$ or $a_2 = a_1/2$ for the present case give the result of Fig. 5. For most asymmetrical modes depicted in Fig. 2, the screech tone did occur close to the peak value of $|\eta_t|/\eta_{t,\max}$. For symmetrical modes the frequencies occur far from the peak, in fact close to the adjacent minimum. This suggests that there is little or no reinforcement by multiple sources, implying the existence of essentially one single acoustic source for the symmetrical mode, located near the first shock wave.

The symmetrical mode may be encouraged by (a) the frequency is high for pressure ratios just above choking, and for these short wavelengths the expected growth rate for the symmetrical mode is comparable to its asymmetrical counterpart, (b) for nozzles with a relatively small AR, the situation can no longer be considered to be two dimensional, tending to the square nozzle (AR = 1) where symmetrical mode dominates at lower-pressure ratio, and (c) because of its constancy of phase all around the nozzle perimeter, the disturbance effectiveness η_d (see Powell, 1953a) of the symmetric feedback is more efficient than the asymmetrical mode. However, the symmetrical mode is found to be very weak presumably because (a) the presence of only one single significant source, (b) the shock structure is very weak at pressures just above choking, and (c) the disturbances do not have a long distance in which to grow.

III. CONCLUSION

The symmetric mode for choked-jet screech has been found at pressure ratios just above choking, for a rectangular jet of AR = 4.8. The symmetric mode appears due to acoustic

feedback to the nozzle exit from the sound source at the first shock wave. It is relatively weak, apparently mainly due to the stream instability having little distance in which to grow before interacting mainly with only the first shock structure, which itself is weak.

Brown, G. B. (**1937**). ''The vortex motion causing edge-tones,'' Proc. Phys. Soc. London **49,** 493–507.

Kozu, T. (**1991**). ''An experimental study of a two-dimensional jet above the critical pressure,'' unpublished report, Department of Mechanical Engineering, University of Houston.

Krothapalli, A. (**1986**). ''The role of screech tones in mixing of an underexpanded rectangular jet,'' J. Sound Vib. **106,** 119–143.

Lin, D. (**1992**). ''Sound produced by the impingement of a choked jet issuing from a rectangular nozzle on a wedge,'' M.S. thesis, Department of Mechanical Engineering, University of Houston.

Mascarenhas, N. (**1992**). ''Subsonic edge-tones,'' M.S. thesis, Department of Mechanical Engineering, University of Houston.

Mattingly, G. E., and Chang, C. C. (**1974**). ''Unstable waves on an axisymmetric jet column,'' J. Fluid Mech. **65,** 541–560.

Powell, A. (**1953a**). ''On edge tones and associated phenomena,'' Acustica **3,** 233–243.

Powell, A. (**1953b**). ''On the noise emanating from a two-dimensional jet above the critical pressure,'' Aero. Quart. **4**, 103–122. (The larger dimension of the nozzle of 0.70 in. was erroneously stated to be 0.07 in.)

Powell, A. (**1962**). ''Nature of the feedback mechanism on some fluid flows producing sound,'' Fourth International Congress on Acoustics, Copenhagen.

Powell, A., Umeda, Y., and Ishii, R. (**1990**). ''The screech of round choked jets, revisited,'' AIAA Paper AIAA-90-3980, AIAA 13th Aeroacoustics Conference.

Raman, G., and Rice, E. J. (**1994**). ''Instability modes excited by natural screech tones in a supersonic rectangular jet,'' Phys. Fluids **6,** 3999–4008.

Rayleigh, J. W. S. (**1945**). *The Theory of Sound* (Dover, New York), Vol. 2.

Tam, C. K. W. (**1995**). ''Supersonic jet noise,'' Annu. Rev. Fluid Mech. **27,** 17–43.

Umeda, Y., Ishii, R., Matsuda, T., Yasuda, A., Sawada, K., and Shima, E. (**1990**). ''Instability of astrophysical jets II,'' Prog. Theor. Phys. **84,** 856–866.

Umeda, Y., and Ishii, R. (**1994**). ''Oscillation modes of underexpanded jets issuing from square and equilateral triangular nozzles,'' J. Acoust. Soc. Am. **96,** 1853–1857.

Comments on "The attenuation and dispersion of sound in water containing multiply interacting air bubbles" [J. Acoust. Soc. Am. 99, 3412–3430 (1996)]

Zhen Ye[a)]
Institute of Ocean Sciences, Sidney, British Columbia, Canada
and Department of Physics, National Central University, Chungli, Taiwan, Republic of China

(Received 9 August 1996; accepted for publication 10 November 1996)

The new dispersion relation for acoustic wave propagation in bubbly media introduced by C. Feuillade [J. Acoust. Soc. Am. **99**, 3412–3430 (1996)] is rederived from a self-consistent approach, and examined through numerical examples. A few concerns are raised with regard to the validity of the assumptions that lead to the new dispersion relation. © *1997 Acoustical Society of America.* [S0001-4966(97)04407-X]

PACS numbers: 43.30.Es, 43.30.Ft, 43.30.Lz [JHM]

Multiple scattering of acoustic waves in water containing many air bubbles has become an increasingly important area of underwater acoustics research. The work of Foldy,[1] Lax,[2] Waterman *et al.*,[3] and Twersky[4] serves as a foundation to the field, followed by the studies including high-order bubble–bubble interactions.[5,6] The study of multiple scattering in bubbly liquids has been further advanced by the recent work of Feuillade[7] who provided a new dispersion relation for wave propagation through bubbly media. The purposes of this communication are: (1) to provide an alternative but equivalent derivation of the new dispersion relation; (2) to examine critically the assumptions that lead to the result; and (3) to inspect some consequences of the result. A reason for rederiving the new dispersion relation is our belief that the present derivation can identify more explicitly the assumptions that underlie the new relation. This facilitates our later discussion of their validity. Through the process, it will become clear that the new dispersion relation is problematic.

Consider a unit plane wave, described by $p_0(\mathbf{r}) = e^{i\mathbf{k}_{\rm inc}\cdot\mathbf{r}}$, propagating through a water containing N air bubbles of the same size located at $\mathbf{r}_i$ $(i=1,2,\ldots,N.)$ Here $\mathbf{k}_{\rm inc}$ is the incident wave vector and $|\mathbf{k}_{\rm inc}|=k=\omega/c$. In the mean field approximation,[1,5] the dispersion relation can be written as

$$\kappa^2 = k^2 + 4\pi n F, \tag{1}$$

where κ is the effective wave number when the bubbles are present, n is the bubble population function (number/m^3), and F is an effective frequency-dependent scattering function of a single bubble. To determine F, a self-consistent method may be invoked. We generically write the scattered wave from the ith bubble as

$$p_s^i(\mathbf{r}) = A_i \frac{e^{ik|\mathbf{r}-\mathbf{r}_i|}}{|\mathbf{r}-\mathbf{r}_i|}. \tag{2}$$

This is valid because we consider low-frequency scattering, for which the isotropic scattering due to the radially pulsating mode dominates. On the other hand, the scattered wave is a linear response to the total incident wave, which includes the direct incident wave and the scattered wave from other bubbles, and therefore can be written alternatively as

$$p_s^i(\mathbf{r}) = f\left(p_0(\mathbf{r}_i) + \sum_{j=1,j\neq i}^{N} A_j \frac{e^{ik|\mathbf{r}_i-\mathbf{r}_j|}}{|\mathbf{r}_i-\mathbf{r}_j|} \right) \frac{e^{ik|\mathbf{r}-\mathbf{r}_i|}}{|\mathbf{r}-\mathbf{r}_i|}, \tag{3}$$

where f is the scattering function of a single bubble,[8]

$$f = \frac{a}{\omega_0^2/\omega^2 - 1 - i\delta},$$

in which ω_0 is the resonance frequency, δ is the damping factor, and a is the bubble radius. Equating Eqs. (2) and (3), we arrive at the following self-consistent equations, which are equivalent to the equation derived originally by Foldy,[1]

$$A_i = f\left(p_0(\mathbf{r}_i) + \sum_{j=1,j\neq i}^{N} A_j \frac{e^{ik|\mathbf{r}_i-\mathbf{r}_j|}}{|\mathbf{r}_i-\mathbf{r}_j|} \right), \quad i=1,2,\ldots,N. \tag{4}$$

In principle, the N unknown coefficients A_i can be solved from these N equations when N is not too large, for which reason we call it few body system. Unfortunately, N is a large number for most bubble systems of interest, which we refer to as many-body systems. For a many-body system, it would be convenient to use diagram methods[6,9] to solve these coupled equations. Once A_i is determined, the effective scattering function F can be obtained, according to the Sommerfeld radiation condition:

$$F = \lim_{r\to\infty} \frac{p_s^i(\mathbf{r})}{p_0(\mathbf{r}_i)} \left(\frac{e^{ik|\mathbf{r}_i-\mathbf{r}_i|}}{|\mathbf{r}-\mathbf{r}_i|} \right)^{-1} = \frac{A_i}{p_0(\mathbf{r}_i)}. \tag{5}$$

Here we have taken the ith bubble as an example. To the lowest-order approximation, A_i equals $f p_0(\mathbf{r}_i)$, and therefore $F=f$. Substituting this into Eq. (1), we are led to the classic result of Foldy,[1]

$$\kappa^2 = k^2 + 4\pi n f. \tag{6}$$

In order to obtain the new dispersion relation of Feuillade, we must make the following crucial assumption from Feuillade [Refer to p. 3418 and Eq. (21) in Ref. 7]:

Assumption I: All the interacting bubbles have the same scattering coefficients, i.e.,

[a)]Electronic mail: zhen@pinger.ios.bc.ca

$$A_j = \cdots = A_i .$$

(Note that the assumption in Feuillade has been expressed in a different but equivalent form.) This assumption immediately leads to

$$A_i = \frac{p_0(\mathbf{r}_i)}{f^{-1} - \sum_{j=1, j\neq i}^{N} (e^{ik|\mathbf{r}_i - \mathbf{r}_j|}/|\mathbf{r}_i - \mathbf{r}_j|)}. \tag{7}$$

Inserting Eq. (7) into Eq. (5) and using Eq. (1), we obtain exactly the dispersion relation of Feuillade,[7]

$$\kappa^2 = k^2 + \frac{4\pi n}{f^{-1} - 4\pi n \int_0^\infty dr\ r e^{ikr}}, \tag{8}$$

in which the summation is replaced with the spatial integral as in Feuillade,[7] as a result of ensemble average, i.e.,

$$\sum_{j=1, j\neq i}^{N} (\,) = 4\pi n \int dr\ r^2 (\,).$$

We note that, in actuality, the integral in the above equation should be carried out from $2a$ to infinity. However, since $ka \ll 1$, the lower limit can be approximated as 0 in the cases under consideration. Furthermore, the wave number k appearing in the first term on the rhs of Eq. (8) represents the wave number in the pure water. In the original derivation, the wave number k in the integrand should also be this wave number. However, in self-consistent methods, the latter wave number is allowed to be incorporated with an imaginary part to denote the attenuation effect, and it may even be substituted by κ for a closed self-consistent calculation as in the well-known Hartree–Fock approximation. In the following, all these cases will be addressed.

The validity of Assumption I deserves careful consideration. As shown by Eq. (7), each A_i carries the phase factor due to the direct incident wave, i.e., $i\mathbf{k}_{\text{inc}} \cdot \mathbf{r}_i$. This would induce a phase difference between two interacting bubbles, i and j say, as $\mathbf{k}_{\text{inc}} \cdot \mathbf{r}_i - \mathbf{k}_{\text{inc}} \cdot \mathbf{r}_j$, which is of the same order of magnitude as the phase factor $ik|\mathbf{r}_i - \mathbf{r}_j|$ in the new dispersion relation. The term involving this latter phase factor is referred to as the radiative effect between bubbles by Feuillade.[7] Therefore, if Assumption I holds, this latter phase factor should not appear in the new dispersion relation. Then the integral in Eq. (8) would lead to an unphysical divergence. In other words, if the phase difference $ik|\mathbf{r}_i - \mathbf{r}_j|$ is to be kept, then the phase difference due to the incident wave should also be retained. Now leaving discussion on the validity of Assumption I for a moment, we inspect the consequences of the new dispersion relation.

The integral in Eq. (8), a Fourier transformation, can be evaluated as $-(1/k^2)e^{2ika} + (2ia/k)e^{2ika}$ for finite a and $k \neq 0$. This can be approximated as $-1/k^2$ for the small bubbles under consideration. (Numerical calculation confirms this.) This result can be obtained mathematically by adding an imaginary infinitesimal to k. Physically, k always has an imaginary part, representing the acoustic absorption as the wave propagates through the medium. We consider the attenuation in the case shown in Fig. 6 of Feuillade as an example. The bubble void fraction is taken as 0.22%, and the bubble radius as 1.77 mm. The attenuation coefficient in dB per unit length is given as usual 8.68 Im(κ).[8] Figure 1 shows

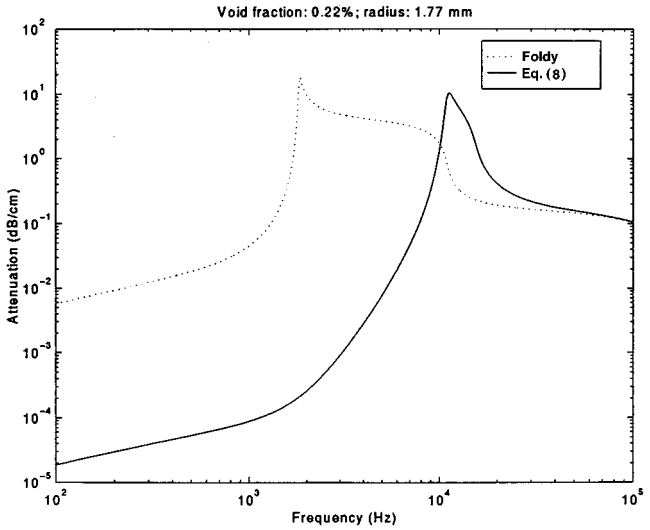


FIG. 1. The attenuation coefficient as a function of frequency from Eq. (8): the bubble void fraction 0.22%, and the bubble radius 1.77 mm.

the attenuation coefficient as a function of frequency using Eq. (8). The significant discrepancy between the classic and new dispersion relations is shown clearly. Feuillade[7] suggested that the wave number in the integral in Eq. (8) be replaced by κ. This gives a self-consistent equation

$$\kappa^2 = k^2 + \frac{4\pi n}{f^{-1} - 4\pi n \int_0^\infty dr\ r e^{i\kappa r}}, \tag{9}$$

which can be rearranged as

$$\kappa^2 = k^2 \left(1 + \frac{4\pi n f}{\kappa^2} \right). \tag{10}$$

The attenuation coefficient from this equation is plotted in Fig. 2 against frequency. Apparently, the result differs from what Feuillade had shown in his Fig. 4. Here we see that the new dispersion relation leads to quite different results from the Foldy formulation. Again we note that the lower limit of

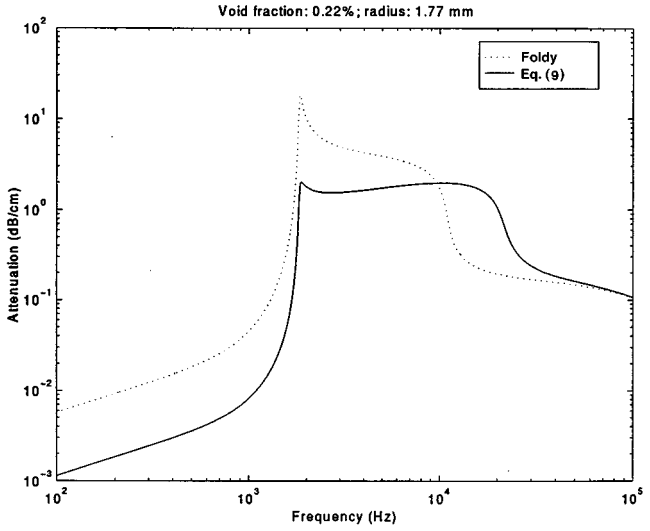


FIG. 2. The attenuation coefficient as a function of frequency from Eq. (9): the bubble void fraction 0.22%, and the bubble radius 1.77 mm.

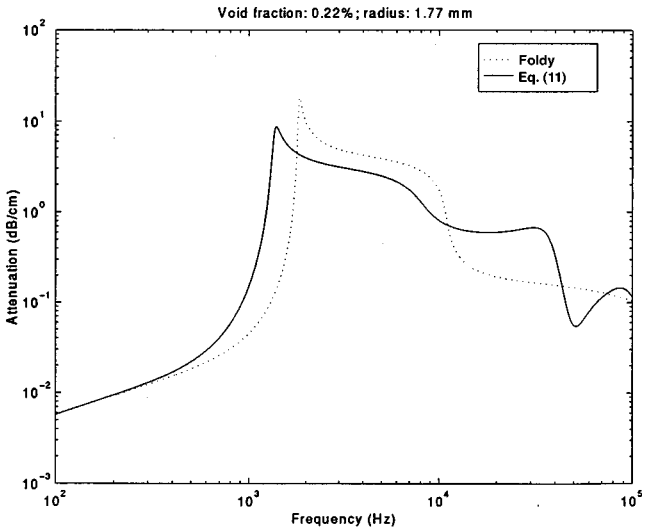


FIG. 3. The attenuation coefficient as a function of frequency from Eq. (11): the bubble void fraction 0.22%, the bubble radius 1.77 mm, and the cutoff range 2.79 cm.

the integral in Eq. (9) is approximated as 0, which has been justified by the numerical computation.

In order to fit the experimental data,[10] Feuillade later introduced a frequency-independent cutoff range R for the integral in Eq. (8).

Assumption II: Effectively the interaction between bubbles is confined within a range R, i.e.,

$$\int_0^{\infty} dr\ re^{ikr} \rightarrow \int_0^{R} dr\ re^{ikr}.$$

The dispersion relation becomes

$$\kappa^2 = k^2 + \frac{4\pi n}{f^{-1} - 4\pi n \int_0^R dr\ re^{ikr}}. \tag{11}$$

For the above case, R is chosen as 2.79 cm, slightly larger than the average distance between two bubbles. Then the new attenuation coefficient is plotted as a function of frequency in Fig. 3. The result is consistent with Fig. 6 of Feuillade,[7] which appears to fit the experimental data well. However, it is useful to examine the implications of introducing a cut-off range R. The reasons can be stated as follows in connection with our concerns about Assumption I.

Feuillade indicated that the cutoff range is a result of the screening effect due to multiple scattering. He wrote ''The high level of attenuation near resonance implies that the scattering cross sections of the individual bubbles are very large in this region. They are so large, in fact, that effectively all of the sound scattered from any bubble interacts with, and is rescattered by, only the few bubbles immediately adjacent to it.'' In Assumption II, the cutoff range characterizes the size of the region referred to by Feuillade. This seems physically consistent with Assumption I, that is, the nearby interacting bubbles approximately have the same phase. However, if the assumptions are valid, the cutoff range R must be related to the attenuation resulting from multiple scattering, and has to be a function of frequency, as we can see in the above figures that the attenuation coefficient decreases dramatically when the wave propagates at a frequency not in the vicinity of resonance. This fact has not been incorporated in the present form of Feuillade's theory. At frequencies where the attenuation is not so large, more remote bubbles can interact with each other. In this case, Assumption I may fail. In addition, Assumption I, combined with II, implies that kR must be much smaller than 1, as also indicated by Feuillade. This condition is satisfied for most cases considered by Feuillade. However, there are cases in which this condition fails. For example, consider the above case, $R=2.79$ cm, for which kR varies from 0.0117 at 10^2 Hz to 1.169 at 10^4 Hz, in the frequency range of interest. When kR is comparable to 1, the consistency of the assumptions is open to question.

In conclusion, although the new dispersion relation seems appealing under certain manipulations, the assumptions that underlie the relation and the manipulations are questionable. The work of Feuillade does indicate that the classic Foldy theory may not be sufficient for data interpretation in some cases. But the proper correction to be included remains an open issue. A correct way of solving the problem may be to use the systematic diagram method for multiple scattering in bubbly liquids.[6,11] We must stress that, as in previous studies, the discussion is limited to the situation in which the many bubble system is infinitely large; otherwise boundary effects may need be considered.

ACKNOWLEDGMENTS

Dr. David Farmer is sincerely thanked for useful advice and encouragement in preparing this manuscript. Dr. D. Chu (WHOI) is thanked for checking the programs. The useful comments from Dr. Li Ding (NRIFE, Japan) help clarify the paper greatly.

[1] L. L. Foldy, ''The multiple scattering of waves,'' Phys. Rev. **67**, 107–119 (1945).
[2] M. Lax, ''Multiple scattering of waves,'' Rev. Mod. Phys. **23**, 287–310 (1951).
[3] P. C. Waterman and R. Truell, ''Multiple scattering of waves,'' J. Math. Phys. (N.Y.) **2**, 512–537 (1961).
[4] V. Twersky, ''On scattering of waves by random distributors,'' J. Math. Phys. **3**, 700–715 (1962).
[5] A. S. Sangani, ''A pairwise interaction theory for determining the linear acoustic properties of dilute bubbly liquids,'' J. Fluid Mech. **232**, 221–284 (1991).
[6] Z. Ye and L. Ding, ''Acoustic dispersion and attenuation relations in bubbly mixture,'' J. Acoust. Soc. Am. **98**, 1629–1636 (1995).
[7] C. Feuillade, ''The attenuation and dispersion of sound in water containing multiply interacting air bubbles,'' J. Acoust. Soc. Am. **99**, 3412–3430 (1996).
[8] C. S. Clay and H. Medwin, *Acoustical Oceanography* (Wiley, New York, 1977), Appendix 6.
[9] U. Frisch, ''Wave propagation in random media,'' in *Probabilities Methods in Applied Mathematics*, edited by A. T. Barucha-Reid (Academic, New York, 1968), Vol. 1, pp. 76–198.
[10] E. Silberman, ''Sound velocity and attenuation in bubbly mixtures measured in standing wave tubes,'' J. Acoust. Soc. Am. **29**, 925–933 (1957).
[11] Z. Ye and L. Ding, ''A study of multiple scattering in bubbly liquids by many body theory,'' Can. J. Phys. **74**, 92–96 (1996).

Response to "Comments on 'The attenuation and dispersion of sound in water containing multiply interacting air bubbles'" [J. Acoust Soc. Am. 102, 1239–1241 (1997)]

C. Feuillade
Naval Research Laboratory, Stennis Space Center, Mississippi 39529-5004

(Received 20 February 1997; accepted for publication 22 March 1997)

A response to the comments made by Zhen Ye in a Letter to the Editor about the paper ''The attenuation and dispersion of sound in water containing multiply interacting air bubbles [J. Acoust. Soc. Am. **99**, 3412–3430 (1996)]'' is presented. It shows that, while Ye has made some useful observations, his comments fail to bring the theoretical assumptions of the original paper into question. [S0001-4966(97)04507-4]

PACS numbers: 43.30.Es, 43.30.Jx, 43.30.Ky, 43.30.Ft [JHM]

INTRODUCTION

I wish to thank Zhen Ye for reading my paper so carefully, and for providing a concise statement of the main assumptions underlying my theory. His alternative derivation of the dispersion relation for acoustic propagation in bubbly water is instructive and helpful. He has demonstrated the unfeasibility of using an iterative self-consistent approach to evaluate the dispersion relation, a result I fully agree with, and has also clarified the high-frequency limitations of using a truncated integral approximation for its evaluation.

My paper: ''The attenuation and dispersion of sound in water containing multiply interacting air bubbles'' [J. Acoust Soc. Am. **99**, 3412–3430 (1996)] (''FP''), is concerned with the correction of inaccuracies which are sometimes introduced by the classic theory of acoustic propagation in bubbly water,[1,2] due to its omission of multiple scattering processes between bubbles. Significant inaccuracies arise, however, only for high volume fractions in the resonance frequency region, and the intention of the theory presented in FP is to remedy the failures of the classic theory in this case.

In the resonance frequency region, and when the volume fraction is high, the acoustic behavior is characterized by two important features: (a) the wavelength is much greater than the bubble radius ($\sim$460 times), and also much greater than the average spacing between any two adjacent bubbles in the medium ($\sim$37 times, for an air volume fraction β=0.22%); (b) the bubbles interact with and scatter sound very strongly, so that the acoustic field is quickly attenuated (for β = 0.22%, Silberman[3] measured attenuations between 3 and 8 dB per average bubble spacing in the resonance region, which corresponds to very short decay length constants, i.e., $\sim$ one bubble spacing).

The theory in FP is based on the recognition that these two characteristics, taken together, imply two physical consequences for the collective response of bubbly water to an acoustic field at resonance frequencies: (a) because of the high attenuation effect, the scattered field from any bubble is quickly absorbed and rescattered by its immediate neighbors (thus it is directly coupled by radiation to typically only a few bubbles adjacent to it, implying that the collective action of the medium is a local phenomenon); (b) because of the long wavelength at resonance, the excitation phase varies very slowly over the small spatial region occupied by the group of radiatively coupled bubbles at each locality and, to a good approximation, may be considered constant.

Ye, in his Comments (''YC''), after presenting an alternative derivation of the dispersion relation for bubbly water given in FP, correctly identifies the essence of these ideas when he makes the following assertions, i.e., that FP utilizes two assumptions:

''*Assumption I: All the interacting bubbles have the same scattering coefficients.*''
''*Assumption II: Effectively the interaction between bubbles is confined within a range R.*''

Having made these assertions, YC gives numerical examples in an attempt to question the validity of the dispersion relation of FP. Now, while ''Assumption I'' correctly represents a theoretical assumption made in FP, ''Assumption II,'' strictly speaking, does not. Although ''Assumption II'' admittedly implies the concept of local action discussed in the previous paragraph, the introduction of a single parameter R to represent the range of bubble interactions is an approximation introduced in FP for data fitting purposes. To facilitate discussion, I will adopt the term ''Assumption IIA'' as a convenient handle to connote the idea of the confinement of the collective action to a small locality around each bubble, but not necessarily within a precise boundary delimited by R.

Assumption I and Assumption IIA are interlinked, such that Assumption I could not be valid if Assumption IIA were not valid at the same time. For example, in a large bubble cloud, bubbles at widely separated locations will experience different driving phases. If these widely separated bubbles were also to strongly interact, then Assumption I would not be valid. However, the attenuation is high in the resonance region in densely populated bubble clouds, and FP contends that these adverse conditions do not coincide.

It is the purpose of FP to address the resonance region where both Assumption I and Assumption IIA hold simultaneously, thus constituting a favorable conjunction of circumstances which makes possible the analytical approach used in FP. Off-resonance, and for sparsely populated clouds, where

the interaction between bubbles and the attenuation is much smaller, the classic theory gives an adequate description of the physical behavior.

While YC has identified the assumptions underlying FP, I will show, in this response, that the three examples used in YC to raise concerns about FP fail to bring the assumptions into question.

I. ASSUMPTION I AND ASSUMPTION IIA

The mathematical means by which Assumption I is incorporated into FP is by resolving the collective motion of bubbly water into the individual behavior of normal modes, and then analyzing the characteristics of the dominant ''symmetric'' mode. In a previous paper,[4] I have already acknowledged and explained the close relationship of this approach to the self-consistent derivation of Twersky[5] and others (the method outlined in YC). One of the advantages of the approach taken in FP is that it shows, interestingly, that Assumption I does not necessarily imply that the bubbles oscillate with the same amplitude, but only with the same phase.

Assumption IIA is introduced implicitly in FP through the factor (e^{-ikr}/r), which describes the spherical spreading of the field scattered from each bubble causing the bubbles to mutually interact and couple together. To ensure the local nature and small interactive range of this scattered radiation (i.e., so that only bubbles within a small region are radiatively coupled), k must necessarily contain an attenuation component to describe the decay of the field as it propagates through neighboring bubbles. When k was introduced and defined [via Eq. (5) in FP], it was specifically allowed to take complex values, to allow for this attenuation effect.

It has become evident to me, since FP appeared in the journal, that I did not make this point clearly enough. Several workers have assumed that I intended k to represent the propagation wave number for pure water (i.e., water without any bubbles present), although k is never generally defined to represent this quantity in FP. Perhaps a different notation would have been preferable, but I think a contextual reading of the theory and data analysis sections of FP could make my original intentions clear.

Another issue needing restatement is that the scattered field which couples bubbles together (introduced in FP via the e^{-ikr}/r factor) incorporates both amplitude *and phase* components, even when the bubbles are driven at the same phase by an external field. This is explained in the discussion of symmetric modes in my previous paper.[4]

II. RESPONSE TO EXAMPLES IN YC

A. Figure 1

Ye claims that Eq. (8) of YC is identical to the dispersion relation [Eq. (30)] of FP. His expression is:

$$\kappa^2 = k^2 + \frac{4\pi n}{f^{-1} - 4\pi n \int_0^\infty r e^{ikr}\, dr}. \tag{YC8}$$

Actually, his expression differs from the dispersion relation of FP in a critically important way. In order to be made equivalent it should be written:

$$\kappa^2 = k_w^2 + \frac{4\pi n}{f^{-1} - 4\pi n \int_0^\infty r e^{ikr}\, dr}. \tag{FP30}$$

Here, k (in the integrand) represents some complex wave number whose imaginary component describes attenuation in bubbly water (as consistent with Assumption IIA). The first term on the rhs contains k_w (a real number), which represents the propagation wave number for pure water. Ye implies via his Eq. (8) that the symbol k is equivalent to k_w. In fact, the only place where k is equated to k_w in FP is when the approximation of a ''region of collective interaction'' is introduced later in the experimental data analysis section, where it is explicitly stated that the wave number for pure water is used to evaluate the integral, together with a reduced upper integration limit R. This will be taken up again later, when Fig. 3 in YC is discussed.

Figure 1 in YC, which Ye obtains via his Eq. (8) and by assuming k is equivalent to k_w, demonstrates an incongruous calculation which fundamentally misrepresents the theory of FP. The integral in the denominator of the dispersion relation cannot be evaluated realistically to an upper integration limit of ∞, using the real wave number k_w in the exponent of e, because this approach does not incorporate the attenuation effect of bubbly water. The physical nature of the problem simply does not justify this calculation.

B. Figure 2

Equations (9) and (10) in YC represent an iterative procedure for evaluating the new dispersion relation presented in FP. Ye uses this procedure for the $\beta = 0.22\%$ case, and obtains Fig. 2 in YC. This differs significantly from Fig. 4 of FP, which used the same method of evaluation. I am indebted to Ye for bringing this to my attention. My original calculations contained a programming error. A corrected figure has been produced (see Fig. A here) and incorporated in an erratum.[6] The iterative method of solution was used only for Fig. 4 in FP, and nothing else.

The iterative method of solution was not pursued in FP because problems were suspected in using the macroscopic properties of bubbly water to describe microscopic processes between individual bubbles, which would make this closed self-consistent approach inappropriate. The purpose of Fig. 4 in FP was not to promote the iterative method, but to indicate the possibility of difficulties arising from it. Much the same purpose is served by Fig. A. The solid curve in Fig. A (produced via the iterative method) predicts levels of attenuation in the resonance peak region which are reduced from those indicated by the classic theory, and give a better fit to the 1.77-mm data points (note the logarithmic scales). This is a welcome and expected feature, consistent with results obtained later in FP via the alternative approximate method which was subsequently adopted. However, the overall shape of the solid curve in Fig. A, and its divergence from the data points just below the resonance frequency, are features which make me strongly disinclined to try a general application of the iterative method.

The inclusion of Fig. 2 in YC presumably represents an attempt to suggest doubt about the assumptions underlying FP; but no explanation is offered as to why the figure should

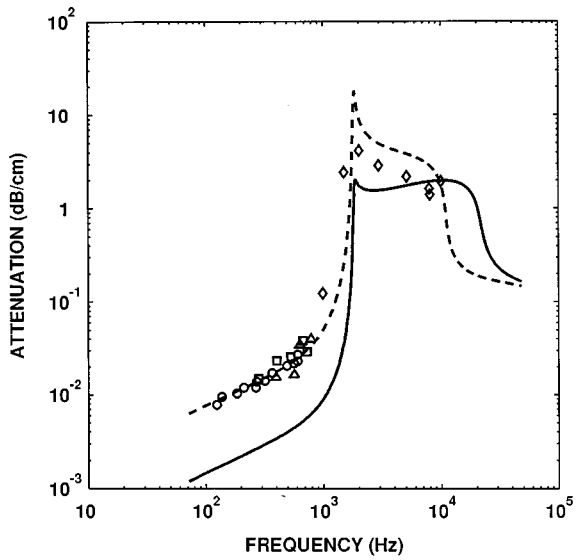


FIG. A. Attenuation coefficient as a function of frequency. Note the logarithmic scales on the axes. The data are taken from Silberman [J. Acoust. Soc. Am. **29**, 925–933 (1957)]. The air volume fraction $\beta = 0.22\%$. The data points are for bubbles of radius: 1.77 mm (''◇''); 1.83 mm (''○''); 2.07 mm (''□''); 2.44 mm (''△''). The dashed line shows the theoretical attenuation coefficient predicted for bubbles of radius 1.77 mm using the classic theory of propagation in bubbly water. The solid line shows the theoretical attenuation coefficient predicted by the dispersion relation of FP for 1.77-mm bubbles, evaluated using an iterative procedure. The shape of the solid line clearly does not follow the data.

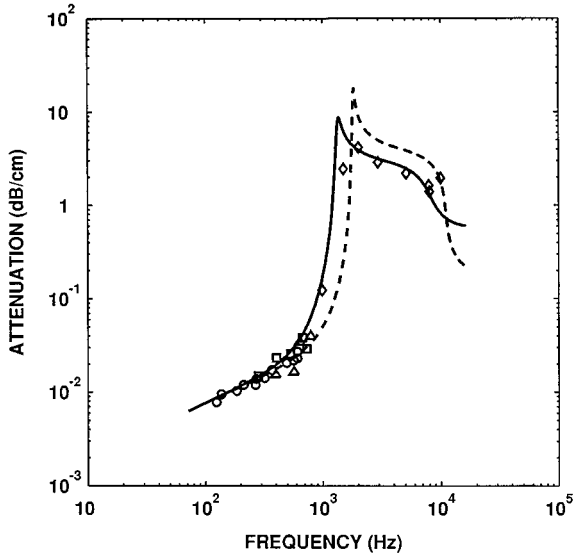


FIG. B. Attenuation coefficient as a function of frequency. The data and the air volume fraction are the same as in Fig. A. The dashed line shows the theoretical attenuation coefficient predicted for bubbles of radius 1.77 mm using the classic theory of propagation in bubbly water. The solid line shows the theoretical attenuation coefficient predicted by the dispersion relation of FP for 1.77 mm bubbles, but by substituting k_w for k and evaluating the integral in the dispersion relation with an upper integration limit $R = 2.79$ cm. This value of R was determined by least-squares fitting to the 1.77-mm points on the plateau above the resonance frequency. For display purposes, both lines have been projected moderately above the highest-frequency data point used in the fitting (i.e., 10 kHz).

actually infer any uncertainty. YC simply comments that the ''...new dispersion relation [i.e., evaluated by the iterative method] leads to quite different results from the Foldy [i.e., classic] formulation...''. This is true, but we can see that, *when data are added* (as in Fig. A), the new dispersion relation actually gives a *better* fit to the points lying in the resonance region (where Assumptions I and IIA are expected to apply) than the classic theory. If anything, this should surely serve to increase, rather than decrease, confidence in the assumptions. While the overall result of applying the iterative method is unsatisfactory, there is no evidence to indicate that this has any bearing on the assumptions underlying FP; and it is more cogently attributable to other causes (see previous paragraph).

C. Figure 3

As mentioned in the introduction, Assumption II (i.e., as originally stated in YC) is not strictly a theoretical assumption underlying FP. It is, in fact, an integral truncation approximation made for data fitting in the resonance frequency region. This is the method used to perform the substantive data analysis in FP, and on which the conclusions of the paper are based. Figure 3 in YC reproduces the curves shown in Fig. 6(a) of FP, except that the calculation has been extended in YC to frequencies far beyond (about 10 times) those necessary to fit the available data. This may be seen by comparing Fig. 3 with Fig. B here. [Figure B is a reproduction of Fig. 6(a) in FP.] The solid curve in Fig. B is a fit to data. It was produced by substituting k_w for k in the dispersion relation of FP (a special procedure used only for this data fitting purpose), and truncating the integral at an upper limit R (instead of ∞), where R was determined by least-squares fitting the solid curve to the 1.77-mm points (''◇'') on the resonance plateau above the peak frequency.

For very high frequencies (corresponding to $kR \gg 1$) this integral truncation approximation cannot be used, because the value of the integral begins to oscillate. The high-frequency end of the solid curve in Fig. 3 of YC shows this. This is an artifact of the approximation, and has no bearing on the theoretical basis of FP. The value of $k_0 R \ll 1$ in Fig. B, where k_0 is the wave number at resonance. This is true for all the data cases analyzed in FP. For most of the points on the resonance plateau in Fig. B the condition $kR < 1$ is also satisfied, and even the highest-frequency data point used for fitting (i.e., at 10 kHz, corresponding to $kR = 1.18$) lies well below the oscillatory region. The remarks made in YC that the theory in FP does not incorporate different values of R for off-resonance points are not relevant, for two reasons: (a) the approximation is not a part of the theory; (b) it was not designed to fit the off-resonance points.

In essence, all that Fig. 3 in YC demonstrates is the break down of an approximation in a region where it was never intended to be used. While other approximations could be devised, in the range where it is applied, this one gives a good fit to the data, and provides useful results by enabling the total number of bubbles involved in the interaction to be estimated.

III. CONCLUSIONS

The alternative derivation of the new dispersion relation for acoustic propagation in bubbly water provided by YC is

helpful and informative. However, while YC has identified the main assumptions underlying FP, the three examples subsequently used to question these assumptions fail to raise any substantive issues. The first example incorporates a calculation which misapplies the theory of FP. The second example detected a simple computational error in FP, which has been corrected in an erratum. The third example is basically a criticism of the single R approximation used in FP for data fitting.

Assumptions I and IIA are important assumptions, and certainly controversial. While they have limitations, in the resonance region there are good arguments for adopting them. In FP they have lead to a meaningful and consistent interpretation of data, and provided answers to some long-standing questions about the importance of multiple scattering processes in oceanic bubble clouds.

[1] L. L. Foldy, ''The multiple scattering of waves. I. General theory of isotropic scattering by randomly distributed scatterers,'' Phys. Rev. **67**, 107–119 (1945); E. L. Carstensen and L. L. Foldy, ''Propagation of sound through a liquid containing bubbles,'' J. Acoust. Soc. Am. **19**, 481–501 (1947).

[2] K. W. Commander and A. Prosperetti, ''Linear pressure waves in bubbly liquids: Comparison between theory and experiments,'' J. Acoust. Soc. Am. **85**, 732–746 (1989).

[3] E. Silberman, ''Sound velocity and attenuation in bubbly mixtures measured in standing wave tubes,'' J. Acoust. Soc. Am. **29**, 925–933 (1957).

[4] C. Feuillade, ''Scattering from collective modes of air bubbles in water and the physical mechanism of superresonances,'' J. Acoust. Soc. Am. **98**, 1178–1190 (1995).

[5] V. Twersky, ''Multiple scattering of waves and optical phenomena,'' J. Opt. Soc. Am. **52**, 145–171 (1962).

[6] C. Feuillade, ''Erratum: 'The attenuation and dispersion of sound in water containing multiply interacting air bubbles' [J. Acoust. Soc. Am. **99**, 3412–3430 (1996)],'' J. Acoust. Soc. Am. **101**, 1746 (1997).

The splitting of dispersion curves for plates fluid-loaded on both sides

X. L. Bao
Department of Mechanical Engineering, Auburn University, Auburn, Alabama 36849

H. Franklin
LAUE, URA CNRS 1373, University of Le Havre, Place Robert Schuman, 76610 Le Havre, France

P. K. Raju
LAUE, URA CNRS 1373, University of Le Havre, Place Robert Schuman, 76610 Le Havre, France and Department of Mechanical Engineering, Auburn University, Auburn, Alabama 36849

H. Überall
LAUE, URA CNRS 1373, University of Le Havre, Place Robert Schuman, 76610 Le Havre, France and Department of Physics, Catholic University of America, Washington, DC 20064

(Received 3 September 1996; accepted for publication 9 March 1997)

For *in vacuo* elastic plates or plates fluid-loaded on one side, it has been previously shown that dispersion curves for both plate-borne (Lamb) type and, in the latter case, also fluid-borne (Scholte–Stoneley) type fail to cross but repel each other, with the waves simultaneously exchanging their physical character. This study is extended here to the case of a plate loaded by two different fluids, where a quantitative calculation of dispersion curves for water on one side, and a lighter or a heavier fluid on the other, is carried out confirming the existence of two different Scholte–Stoneley waves. © *1997 Acoustical Society of America.* [S0001-4966(97)01207-1]

PACS numbers: 43.40.Dx [CBB]

INTRODUCTION

In two publications,[1,2] a study of plate waves was performed and their phase-velocity versus frequency dispersion curves were obtained. Reference 1 was mainly concerned with the Lamb waves on a plate in vacuum, and demonstrated that the dispersion curves for the two families of symmetric (S_i) and antisymmetric (A_i), $i=1,2,\ldots$ waves repeatedly approached but failed to cross each other, while exchanging their physical character (predominantly compressional or predominantly shear) across the point of repulsion. This reference, and especially Ref. 2, also considered the case of a plate fluid-loaded on one side (aluminum plate in water in Ref. 1, schematic solid–fluid in Ref. 2). The presence of the fluid added a fluid-borne Scholte–Stoneley wave (termed A wave) to the plate-borne Lamb waves, and a similar repulsion phenomenon between the lowest Lamb wave (A_0) dispersion curve and the A-wave dispersion curve takes place, shown for an aluminum plate in water in Fig. 1. The light-dotted curve in Fig. 1 labeled A_0-V/V is the A_0-wave dispersion curve on a free plate, $A(w)$ that of the A wave, and A_0-W/V that of the A_0 wave on a plate water-loaded on one side. The latter two curves approach each other around $fd=0.3$ MHz×mm (f=frequency, d=plate thickness) but then repel each other. Physically, the A_0 wave above $fd=0.3$ is plate-borne and the A wave water-borne (asymptoting to the water-speed value $c_w/c_T=0.486$), while below $fd=0.3$ the opposite is the case. This change of physical character was demonstrated in Ref. 1 by a limiting procedure; mathematical proof for the analogous case of the energy of molecular levels is given in Ref. 3. In both physical situations, the repulsion phenomenon is caused by coupling: for the Lamb waves, via coupling of compressional and shear vibrations through the boundary conditions on the plate faces; for the A_0 and the A wave, via coupling by the fluid-loading. Note that this case has been considered in the pioneering work of Grabowska[4] and Talmant[5] (without any discussion of the physical character of the waves).

I. PLATE WITH FLUID-LOADING ON BOTH SIDES

An early study by Osborne and Hart[6] dealt with a special case of the plate fluid-loaded on both sides, namely, one loaded by the same fluid on both sides. In this case, a second Scholte–Stoneley wave is found in addition to the A wave. Due to the symmetry of this case, the A wave which is similar to that in Fig. 1 is now a purely antisymmetric vibration (as are the A_i Lamb-type waves), deserving the label A. The second Scholte–Stoneley wave is a symmetric vibration (as are the S_i Lamb-type waves), and its dispersion curve is close to a nondispersive horizontal line at the height of the sound speed in the ambient fluid.

In Ref. 2, dispersion curves were shown for a plate loaded by different fluids on the two sides. This was calculated schematically, for no specific elastic material or kinds of fluids. Considering two concrete examples, we evaluate here (using plate equations given by Brekhovskikh[7] or Viktorov[8]) the corresponding dispersion curves for the case of an aluminum plate loaded by (a) water and alcohol, a fluid lighter than water, and (b) water and glycerine, a fluid heavier than water. We obtain the shapes of the A- and S-wave dispersion curves (retaining this nomenclature although the symmetry is now lost), and discuss here the repulsion phenomena that appear.

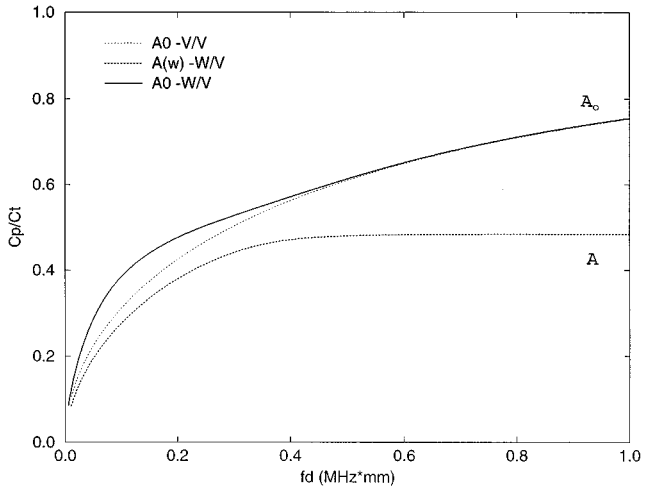


FIG. 1. Phase velocity (c_p) dispersion curves of lowest-order waves on an aluminum plate in vacuum (A_0-V/V), and with one-sided water-loading (aluminum: compressional bulk-wave speed $c_l=6350$ m/s, shear speed $c_t=3050$ m/s, density $\rho=2.7$ g/cm^3; water: $c_w=1483$ m/s, $\rho_w=1$ g/cm^3). Note c_p is normalized by c_t.

Figure 2 presents the dispersion curves for case (a), also showing the free-plate A_0-wave dispersion curve (as a fine line) for comparison. Above $fd=0.3$ MHz×mm we may identify the heavy curve with the A_0 Lamb-type wave, and the lowest, dash-dot curve which asymptotes to the alcohol sound speed ($c_a/c_t=0.393$) with the A wave propagating mainly in alcohol. The dashed curve, asymptoting to the sound speed in water ($c_w/c_t=0.486$) and showing a two-valuedness below point M, propagates mainly in water and may be identified with the S wave. The reason for this is, as was shown by our calculations where c_a was stepwise taken larger until it approached c_w, that the A and the S curves approached each other in the asymptotic region ($fd\geqslant 0.4$), but that in the low-frequency region ($fd\leqslant 0.3$) the A curve always tended to zero while the (dashed) S curve, taken together with the upper branch of its loop, tended toward a

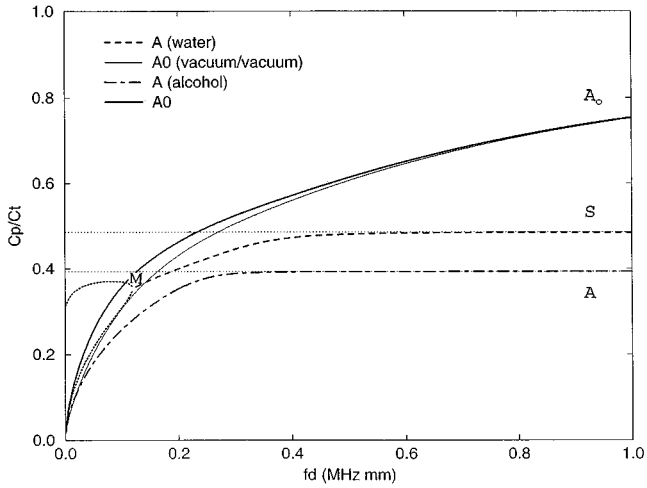


FIG. 2. Phase velocity dispersion curves for lowest-order waves on an aluminum plate in vacuum (light curve), and with water loading on one side and alcohol loading on the other ($c_a=1200$ m/s, $\rho_a=0.8$ g/cm^3).

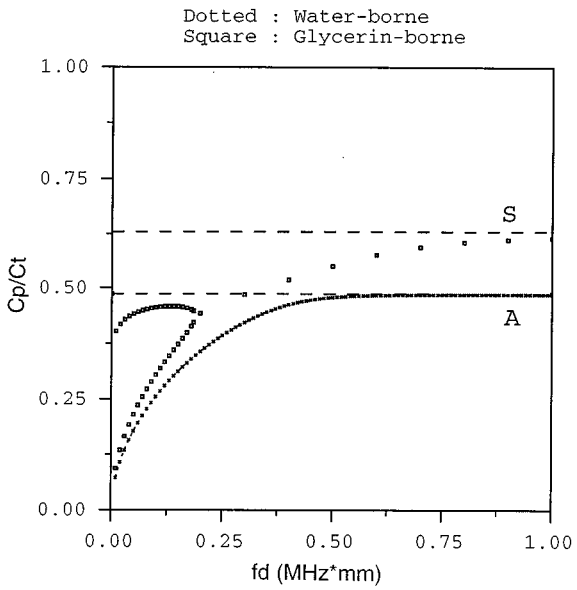


FIG. 3. Phase velocity dispersion curves for lowest-order waves on an aluminum plate with water-loading on one side and glycerine loading on the other ($c_g=1920$ m/s, $\rho_g=1.26$ g/cm^3).

more and more horizontal curve, exactly becoming the S-wave curve in the Osborne and Hart limit.

It should be noted that the double-valued branching of the S-wave dispersion curve is not specific to the present case. Loops of this kind have been observed for, e.g., a steel plate with one-sided fluid-loading.[9]

For case (b), an aluminum plate with water and glycerine on its two sides, Fig. 3 shows the dispersion curves analogous to those of Fig. 2. The A_0 curve is not entered in this figure. We interpret the glycerine-borne wave (asymptoting $c_g/c_t=0.630$) as the S wave (note again its partially double-valued dispersion curve), and the water-borne wave as the A wave. Note that additional studies of the water–glycerine case are contained in the unpublished thesis of Franklin.[10]

II. DISCUSSION

The repulsion phenomenon of dispersion curves can be observed in Figs. 2 and 3 as in the preceding case of one-sided fluid-loading. The reason for this is the coupling of the various plate- and fluid-borne wave types by fluid-loading. As we identified above, in the $fd>0.4$ MHz×mm region, the uppermost dispersion curve belongs to the A_0 wave (shown in Fig. 2 but not in Fig. 3) which is the lowest-order ''antisymmetric'' Lamb wave. The middle dispersion curve belongs to the S-type Scholte–Stoneley wave (mainly propagating in the fluid with faster sound velocity), the lowest to the A wave (mainly propagating in the fluid with slower sound velocity). Passing through the convergence point (at $fd\sim 0.2$ MHz×mm) from above, and taking the one-sided fluid-loading case of Fig. 1 as a guide, it may be asserted that, as before, the A_0 and A waves, being coupled, exchange their characters so that below $fd\sim 0.2$ MHz×mm the A_0 wave (heavy solid curve in Fig. 2) is the ''antisymmetric'' fluid-borne wave and the A wave (dash-dot curve) is the

plate-borne wave. The intermediate S wave (dotted), however, appears not to participate in the coupling, while the lower branch of its loop, at this point, still lacks any detailed physical explanation. It forms with the upper branch of its loop a continuous, almost constant curve resembling the S wave of the Osborne–Hart case, with which it intersects the A_0 curve rather than showing a repulsion phenomenon.

It is seen that the dispersion curves of the lowest-order waves on a plate differently fluid-loaded on both sides form a natural extension of those on a one-sidedly fluid-loaded plate (the Grabowska case), as well as of the A_0, A, and S waves on a plate loaded on two sides by the same fluid (the Osborne–Hart case). We feel, however, that with an awareness of the dispersion curve repulsion phenomenon (especially clearly demonstrated for the Lamb waves in Ref. 1), and of the associated exchange of physical character of the waves as the frequency passes through the convergence point, a physical understanding of the plate- and fluid-borne waves has been gained which a calculation of the dispersion curves alone would not have furnished.

It may be noted that solid- and fluid-borne waves have recently been investigated on a doubly fluid-loaded cylindrical shell, demonstrating the existence of the second Scholte–Stoneley wave and its deviations from the plate case due to the curvature. This study was based on experimental measurements of the scattering of sound pulses and the analysis thereof,[11] rather than on analytical calculations as in the present paper. However, an analytical study of shell scattering and the corresponding shell waves has also been completed at this time.[12] Further shell wave cases have been discussed earlier.[13]

ACKNOWLEDGMENTS

Two of the authors (P. K. R. and H. Ü.) thank Professor J. Ripoche, Director of the Laboratoire d'Acoustique Ultrasonore et d'Electronique, University of Le Havre, France and his colleagues for their hospitality during their stay at that laboratory in the Spring of 1996.

[1] H. Überall, B. Hosten, M. Deschamps, and A. Gérard, ''Repulsion of phase-velocity dispersion curves and the nature of plate vibrations,'' J. Acoust. Soc. Am. **96**, 908–916 (1994).
[2] J. Dickey, G. Maidanik, and H. Überall, ''The splitting of dispersion curves for the fluid-loaded plate,'' J. Acoust. Soc. Am. **98**, 2365–2367 (1995).
[3] See, e.g., P. H. E. Meijer and E. Bauer, *Group Theory* (North-Holland, Amsterdam, 1965), p. 52.
[4] A. Grabovska, ''Propagation of elastic wave in solid layer–liquid system,'' Arch. Acoust. **4**, 57–64 (1979).
[5] M. Talmant, ''Rétrodiffusion d'une impulsion ultrasonore brève par une coque cylindrique à paroi mince,'' Ph.D. thesis, University of Paris VII, 1987; see also M. Talmant and G. Quentin, ''Backscattering of a short ultrasonic pulse from thin cylindrical shells,'' J. Appl. Phys. **63**, 1857–1863 (1988); G. Quentin and M. Talmant, ''The plane plate model applied to the scattering of the ultrasonic waves from cylindrical shells,'' in *Elastic Wave Propagation*, Proceedings of the 2nd IUTAM-IUPAP Symposium on Elastic Wave Propagation, Galway, Ireland, March 20–25, 1988, edited by M. F. McCarthy and M. A. Hayes (Elsevier Science, North-Holland, Amsterdam, 1989).
[6] M. F. M. Osborne and S. D. Hart, ''Transmission, reflection, and guiding of an exponential pulse by a steel plate in water. I. Theory,'' J. Acoust. Soc. Am. **17**, 1–18 (1945); ''II. Experiment,'' J. Acoust. Soc. Am. **18**, 170–184 (1946).
[7] L. M. Brekhovshikh, *Waves in Layered Media* (Academic, New York, 1960).
[8] A. Viktorov, *Rayleigh and Lamb Waves* (Plenum, New York, 1967).
[9] A. Freedman and O. Poncelet (private communication).
[10] Hervé Franklin, ''Propagation des ondes d'interface dans les milieux à plusieurs couches. Application à certains modèles de fonds marins,'' Ph.D. thesis, University of Paris VI, 1991.
[11] J. P. Sessarego, J. Sagéloli, C. Gazhanes, and H. Überall, ''Two Scholte-Stoneley waves on doubly fluid-loaded plates and shells,'' J. Acoust. Soc. Am. **101**, 135–142 (1997).
[12] X. L. Bao, P. K. Raju, and H. Überall, ''Circumferential waves on an immersed, fluid filled elastic cylindrical shell,'' preprint.
[13] H. Überall, A. Gérard, A. Guran, J. Duclos, M. El H. Khelil, X. L. Bao, and P. K. Raju, ''Acoustic scattering resonances: Relation to external and internal surface waves,'' Appl. Mech. Rev. **49**, S63–S71 (1996).

Erratum: "Resonant scattering of acoustic waves by ellipsoid air bubbles in liquids" [J. Acoust. Soc. Am. 101, 681–685 (1997)]

Z. Ye
Department of Physics, National Central University, Chungli, Taiwan, Republic of China

(Received 4 June 1997; accepted for publication 4 June 1997)

[S0001-4966(97)06608-3]

PACS numbers: 43.20.Fn, 43.30.Ft, 43.20.Bi, 43.10.Vx [DWM]

E. Wang is added as a co-author of the paper.

TECHNICAL NOTES AND RESEARCH BRIEFS

Paul B. Ostergaard
10 Glenwood Way, West Caldwell, New Jersey 07006

Editor's Note: *Original contributions to the Technical Notes and Research Briefs section are always welcome. Manuscripts should be double-spaced, and ordinarily not longer than about 1500 words. There are no publication charges, and consequently, no free reprints; however, reprints may be purchased at the usual prices.*

Advanced-degree dissertations in acoustics

Editor's note: Abstracts of Doctoral and Master's theses will be welcomed at all times. Please note that they must be double spaced, limited to 200 words, must include the appropriate PACS classification numbers, and formatted as shown below (don't make the editor retype them, please!). The address for obtaining a copy of the thesis is helpful. Please submit two copies.

The measured upstream impedance for clarinet performance and its role in sound production [43.75.Ef, 43.75.Yy, 43.58.Bh]—Teresa Delaine Wilson, *School of Music, University of Washington, Box 353450, Seattle, WA 98195, December 1996 (Ph.D.).* The resonances of the performer's airways, including the mouth, tongue, throat, and lungs, were measured to test the theory that the performer tunes airway resonances to harmonic frequencies of a tone and were determined from the upstream (airway) impedance, which was measured directly using a one-microphone method and indirectly using the linear continuity of flow equation, the measured downstream (instrument) impedance, and simultaneous measurements of the upstream and downstream pressures. The indirect method was verified for single, normal tones and was used to examine airway resonances during performance. In musical excerpts, the airways were usually tuned to the first or second harmonic. Resonances were slightly stronger at softer dynamic levels, consistent with a decreased glottal opening. A ''closed throat'' had stronger, broader resonances than an ''open throat,'' which produced the better tone quality. Both the airways and the reed resonance may be involved when playing clarion tones without the register key. For pitchbend, a strong airway resonance was at the fundamental frequency. For multiphonics, an airway resonance frequency was a simple linear combination of the instrument resonance frequencies, and adjustment of airway resonance frequencies produced different multiphonics with the same fingering.

Thesis advisor: Douglas H. Keefe.

Acoustic power and intensity based active noise control [43.20.Bi, 43.50.Ki, 43.60.Bf]—Seong-Woo Kang, *Department of Mechanical Engineering, KAIST (Korea Advanced Institute of Science and Technology), Taejon, Korea, March 1996 (Ph.D.).* This thesis attempts to find useful active noise control strategies which consider sound power, sound intensity, and sound energy as their control objectives, based on the acoustic energetics. Acoustic power and intensity based active noise control strategies are suggested to achieve noise reduction in the global space which one desires to control. The active control strategy which regards the acoustic radiation power of sources as a cost function is reformulated in the frequency domain to find the possible conditions for reducing global noise in an enclosed space. The causality in the active source power control system, which is not revealed by frequency-domain analysis, is analyzed by applying classical Wiener filtering technique in time domain. The active control strategy based on the mean sound intensity, that is, the net acoustic power flow, is studied to achieve global noise control by blockading a noise transmission path from one acoustic domain to another one. Two representative cases: interior global noise control in a heavy industrial equipment cabin and exterior global noise control out of a finite-length duct exit, are dealt with by various experiments, to verify the potentials of power control and intensity control respectively.

Thesis advisor: Yang-Hann Kim.

Thesis may be obtained from Center for Noise and Vibration Control, Department of Mechanical Engineering, KAIST, Science Town, Taejon, 305-701, Korea. Author is now at the Core Technology Research Center, Corporate Technical Operations, Samsung Electronics Co., 416, Maetan-3Dong, Paldal-Gu, Suwon, Kyungki-Do, 442-742, Korea.

Control of acoustic cavitation with application to lithotripsy [43.25.Yw, 43.80.Sh]—Michael R. Bailey,*Department of Mechanical Engineering, The University of Texas at Austin, Austin, TX 78712-1063, May 1997 (Ph.D.).*Control of acoustic cavitation (sound-induced bubble activity) is the subject of this dissertation. Application is to clinical lithotripsy where cavitation contributes to kidney stone comminution and tissue damage. An electrical spark at the near focus of an underwater ellipsoidal reflector was the acoustical source, as in the Dornier HM3 lithotripter. Experiments were done with rigid reflectors, pressure release reflectors, and pairs of reflectors sharing a common focus and a controlled delay between sparks. Since a bubble hit by a single shock pulse can grow profoundly and then collapse violently, the hypothesis was that a second pulse timed to arrive during the collapse phase would intensify the ultimate collapse. Experiments and numerical calculations confirmed the hypothesis. Pitted by bubble collapses, aluminum foil placed along the reflector axis recorded the spatial cavitation field. Tempered collapse was also discovered: A sufficiently short delay between two pulses or a change in the order of the positive and negative phases of a single pulse stifled bubble growth. Early collapse was detected acoustically. Computations of bubble radius and collapse pressure reinforced the observations. Finally, two reflectors facing each other produced highly localized, intense cavitation. [Work supported by ONR, NIH#DK43881, and ARL IR&D.]

Thesis advisor: David T. Blackstock.

Adapted version of dissertation available as Technical Report ARL-TR-97-1 (3 March 1997), Applied Research Laboratories, The University of Texas at Austin, P.O. Box 8021, Austin, TX 78713-8021. Technical report also available from DTIC-OCC, Defense Technical Information Center, 8725 John J. Kingman Rd., Suite 0944, Ft. Belvoir, VA 22060-6218.

General elastic anisotropy: intrinsic or caused by damage [43.35.Cg, 43.35.Zc]—Christophe Aristégui,*Laboratoire de Mécanique Physique, URA CNRS 867, Université Bordeaux I, 351 Cours de la Libération, 33405 Talence Cedex, France, March 1997 (Ph.D.).*The ultrasonic characterization methods were limited to orthotropic elastic solids. They are generalized to identify the anisotropy induced by damage on composites materials which lose this symmetry during tensile test. In order to treat this problem without the limiting assumption that the material keeps its orthotropic symmetry in an *a priori* known frame, the total elastic anisotropy is tackled by using phase velocity measurements of bulk wave propagating inside a plate immersed in water. The accurate identification of the 21 components of the most general stiffness tensor has been demonstrated. The difficulty of determination increases as the anisotropy decreases. If the material possesses no known plane of symmetry, the research of the higher symmetry frame using only the wave speed measurements allows one to determine the natural symmetry of the medium and the independent components of the stiffness tensor. Finally, a fully anisotropic degradation of a ceramic matrix composite subjected to a tensile stress in a nonprincipal direction, is emphasized. The rotation of the elasticity principal frame, caused by the load, is shown and the localization of the principal frame of the damage tensor, defined by the change of the stiffness tensor, gives an estimation of the microcrack orientation.

Thesis advisor: Prof. Stéphane Baste.

 0001-4966/97/102(2)/1250/1/$10.00