## Jerry H. Ginsberg

Acoustics
A Textbook for
Engineers
and Physicists
Volume II: Applications

## Acoustics-A Textbook for Engineers and Physicists



(s)


Jerry H. Ginsberg

# Acoustics-A Textbook for Engineers and Physicists 

Volume II: Applications

Jerry H. Ginsberg<br>G. W. Woodruff School of Mechanical Engineering<br>Georgia Institute of Technology<br>Dunwoody, GA<br>USA

ISBN 978-3-319-56846-1
ISBN 978-3-319-56847-8 (eBook)
DOI 10.1007/978-3-319-56847-8
Library of Congress Control Number: 2017937706
© Springer International Publishing AG 2018
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper
This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

## The ASA Press

The ASA Press imprint represents a collaboration between the Acoustical Society of America and Springer dedicated to encouraging the publication of important new books in acoustics. Published titles are intended to reflect the full range of research in acoustics. ASA Press books can include all types of books published by Springer and may appear in any appropriate Springer book series.

Editorial Board
Mark F. Hamilton (Chair), University of Texas at Austin James Cottingham, Coe College
Diana Deutsch, University of California, San Diego
Timothy F. Duda, Woods Hole Oceanographic Institution
Robin Glosemeyer Petrone, Threshold Acoustics
William M. Hartmann (Ex Officio), Michigan State University
James F. Lynch (Ex Officio), Woods Hole Oceanographic Institution
Philip L. Marston, Washington State University
Arthur N. Popper (Ex Officio), University of Maryland
Martin Siderius, Portland State University
Andrea M. Simmons, Brown University
Ning Xiang, Rensselaer Polytechnic Institute
William Yost, Arizona State University

## The Acoustical Society of America

On December 27, 1928, a group of scientists and engineers met at Bell Telephone Laboratories in New York City to discuss organizing a society dedicated to the field of acoustics. Plans developed rapidly and the Acoustical Society of America (ASA) held its first meeting on May 10-11, 1929, with a charter membership of about 450. Today, ASA has a worldwide membership of 7000.

The scope of this new society incorporated a broad range of technical areas that continues to be reflected in ASA's present-day endeavors. Today, ASA serves the interests of its members and the acoustics community in all branches of acoustics, both theoretical and applied. To achieve this goal, ASA has established technical committees charged with keeping abreast of the developments and needs of membership in specialized fields as well as identifying new ones as they develop.

The technical committees include acoustical oceanography, animal bioacoustics, architectural acoustics, biomedical acoustics, engineering acoustics, musical acoustics, noise, physical acoustics, psychological and physiological acoustics, signal processing in acoustics, speech communication, structural acoustics and vibration, and underwater acoustics. This diversity is one of the Society's unique and strongest assets since it so strongly fosters and encourages cross-disciplinary learning, collaboration, and interactions.

ASA publications and meetings incorporate the diversity of these technical committees. In particular, publications play a major role in the Society. The Journal of the Acoustical Society of America (JASA) includes contributed papers and patent reviews. JASA Express Letters (JASA-EL) and Proceedings of Meetings on Acoustics (POMA) are online, open-access publications, offering rapid publication. Acoustics Today, published quarterly, is a popular open-access magazine. Other key features of ASA's publishing program include books, reprints of classic acoustics texts, and videos.

ASA's biannual meetings offer opportunities for attendees to share information, with strong support throughout the career continuum, from students to retirees. Meetings incorporate many opportunities for professional and social interactions, and attendees find the personal contacts a rewarding experience. These experiences result in building a robust network of fellow scientists and engineers, many of whom became lifelong friends and colleagues.

From the Society's inception, members recognized the importance of developing acoustical standards with a focus on terminology, measurement procedures, and criteria for determining the effects of noise and vibration. The ASA Standards Program serves as the Secretariat for four American National Standards Institute Committees and provides administrative support for several international standards committees.

Throughout its history to present day, ASA's strength resides in attracting the interest and commitment of scholars devoted to promoting the knowledge and practical applications of acoustics. The unselfish activity of these individuals in the development of the Society is largely responsible for ASA's growth and present stature.

To Leah Morgan, Elizabeth Rachel, and Abigail Rose, my grandchildren. Each is talented, each is beautiful, each is unique, each is amazing. I love them.

## Preface

## The Basic Concept

Because you are reading this book, there is a strong likelihood that you are familiar with its companion, Volume 1. If that is so, then you are well acquainted with my core philosophy of instruction. It is evident that students should be prepared to address technical issues in the future. An essential component in doing so is the instructor, but even the most talented instructor needs support. Fulfilling that need has been the motivation for all my prior books, but even more so it is my objective for these books. To meet the present and future needs of students and instructors, treatments of various principles and concepts feature extensive explanations of the motivation and organization of the derivation, as well as thorough descriptions of the steps that are implemented. Some of the derivations and explanations I believe are unique to this book. In most cases, the derived principles are accompanied by discussions of their physical meaning.

Examples are numerous. All are my own creation. Indeed, the thought of creating a suitable set dissuaded me for a long time from beginning to write Acoustics-A Textbook for Engineers and Physicists, even though I believed that there was a strong need for a text like it. Most examples are more than simple applications of derived formulas. I selected many of these examples not only to illustrate the associated theory, but also to be simplified versions of issues the student might encounter in practice. Equally important to their selection was whether treating the results as the basis for small case studies would enlighten the student about the phenomena associated with that theory. I also used the examples as a vehicle to bring to the fore the fact that in many situations, alternative formulations of an analysis might be viable. I endeavored to use the examples as a way of assisting the student to recognize when these alternatives exist, and to assist them to recognize in other situations which one is best. At the same time, I endeavored to recognize the imperative that the examples be cognizant of the capabilities that can be expected of a student who is being exposed to acoustics for the first time.

Each solution explains why the example is important, why the solution proceeds as it does, how to perform unfamiliar operations, what can be learned from the results about fundamental behaviors, and why the qualitative aspects of the results are consistent with the underlying fundamental principles. Some examples analyze systems by more than one method. This serves to enhance the student's fundamental understanding of the underlying physical processes, as well as enhancing the ability to make the appropriate line of attack when confronted with a new situation. The advent and wide availability of computational software is exploited to lend greater realism to some examples. When the usage of software entails any potentially problematic aspects, especially concerning algorithms and their implementation, those issues are addressed explicitly, sometimes with program fragments. In recognition of the importance of computations, and to help students concentrate on the acoustical aspects when they solve homework exercises, the MATLAB code used to solve the examples is available for download from the Springer server.

At its inception, Acoustics-A Textbook for Engineers and Physicists was a single volume. The sequence of topics for the early chapters was a slightly modified version of the courses I taught at Georgia Tech. I wrote the later chapters to fill the spectrum of subjects that I consider to constitute the core concepts and techniques a student in physics or engineering is likely to encounter. Space limitations, as well as my desire that this should be a textbook, rather than a monograph, dictated the scope in the later chapters. For this reason, they were written with the notion that they should expose the student to the fundamental phenomena and provide the fundamental tools required to study and research these phenomena. No chapter attempts to bring the student up to the current state of the art in that subject.

Some instructor's might consider that some chapters delve too deeply into the subject. I do not agree with this sentiment. Some parts address questions students have asked me, and the development in other parts serves to motivate students by demonstrating interesting and enlightening phenomena. Nevertheless, if one does not desire to take advantage of the depth of treatment, the first few sections of each chapter should be adequate to proceed to the next chapter.

After I completed this book, it was evident that it would be quite large. This caused me to recall my days long ago when I was a student, and disliked carrying very large textbooks. Concurrently, when I surveyed the Table of Contents, I realized that the manuscript divided naturally into two parts. The first six chapters constitute a foundation that I consider requisite knowledge for anyone active in the physical aspects of acoustics. My examination also led me to the realization that the nature of the last seven chapters was different. Few individuals are equally familiar with all of them. This certainly was true in my case. Filling in the gaps in my knowledge, particularly in Chap. 11 on geometrical acoustics, was quite enriching and fulfilling. I wanted to provide that same sense to instructors who might not be familiar with the subject of one or more of the later chapters. I also wanted to prepare students for further study in each of those topics. I foresee using Volume 2 either in a survey course covering selected or all chapters, or else to begin specialized courses on the subject of a single chapter. Both usages would have been followed when I taught acoustics at Georgia Tech. The Acoustics II course covered

Chap. 7 on radiation from objects in free space, Chap. 8 on radiation from vibrating regions in a baffle, the beginning portion of Chap. 9 on waveguides, and the beginning portion of Chap. 10 on enclosures. There also were specialized courses on propagation on the ocean and atmosphere, nonlinear acoustics, and structural acoustics. The associated chapters would not have been adequate for those specialized courses, but I believe they would provide a strong foundation with which to initiate such courses.

## Technical Content

The chapters to be found in this volume are sequenced according to what I believe is sensible in terms of the level of sophistication and analytical difficulty. The highlights of each chapter were discussed in the Preface to Volume I, but it is appropriate to summarize the scope of each chapter. Radiation from vibrating objects is the subject of Chap. 7. Configurations that are addressed are spheres and hemispheres, infinitely long cylinders, and three numerical techniques that exemplify the formulations in current use. Radiation from a piston in a baffle is encountered in a diverse set of applications, so Chap. 8 is devoted to a thorough exploration of farfield and nearfield properties. Two examples explore radiation resulting from square wave motion. The difference between these examples is whether the analysis is formulated in the frequency or time domains. The reader might find the results to be quite interesting.

The field within a waveguide is the subject of Chap. 9. First to be studied is the Webster horn equation for one-dimensional waveguides. The investigation of two-dimensional waveguides introduces the concept of modal analysis. It concludes with an investigation of coupling of the acoustic field with elastic walls that are described by plate theory. The basic theory for three-dimensional waveguides is developed. It is applied to waveguides whose cross section is rectangular and then cylindrical. Chapter 10 is devoted to the sound field within an enclosed region. It begins by using the waveguide representation to explain the alternative descriptions of the field as a set of waves that propagate in multiple directions, or as a set of cavity modes that are standing waves. Both analytical descriptions are developed, with emphasis on the situations where each is best employed. An interesting example uses an infinite series of cavity modes to describe the field within a two-dimensional rectangular enclosure due to a point source. After the analysis, an alternative using the method of images highlights the notion that the selection of an optimal analytical approach sometimes requires consideration of what the objective is. The closure of this chapter describes the Rayleigh-Ritz method and a formulation according to Dowell's approximation. The former is often included in standard texts, but the latter is relatively recent.

Chapter 11 is devoted to geometrical acoustics. Mean flow effects are excluded in order to emphasize how high-frequency rays and wavefronts may be determined. A vertically stratified fluid, which is a fundamental model for ocean acoustics,
is examined first. After that, the ray tracing equations for media whose properties depend arbitrarily on location are developed. Algorithms for solving the governing equations in each case are formulated. Of particular interest is the example of the passage of a plane wave through fluid that is cylindrically heterogeneous. Demonstrations of a caustic and folding of the wavefront beyond the caustic are quite captivating. The chapter closes with a presentation of Fermat's principle. The calculus of variations is developed there.

Scattering from bodies surrounded by a fluid is the subject of Chap. 12. The Born approximation for heterogenous media is derived. Rayleigh scattering and its relation to the Born approximation are the next topic. The metrics commonly sought from a scattering study, such as target strength, are discussed. After that, Kirchhoff scattering theory and its relation to geometrical acoustics are explored. The chapter closes with the application of spherical harmonics to analyze scattering from a rigid sphere and a spherical shell. These studies shed light on the transition from Rayleigh to Kirchhoff scattering, as well as the fundamental importance of fluid loading relative to elasticity.

Chapter 13 closes the textbook with an exploration of nonlinear acoustic analyses and phenomena. The bulk of the chapter is devoted to simple plane waves described by the Riemann solution. Techniques for evaluating it are discussed and used to examine harmonic generation and depletion, followed by the propagation of shocks. The Rankine-Hugoniot relations for weak shock are derived and shown with simple mathematics to lead to the phenomena of old age and acoustic saturation. A nonlinear wave equation is derived as the basis for study of multidimensional nonlinear waves. Its first usage leads to differential equations governing the position dependence of Fourier series coefficients for a plane wave in a dissipative fluid. Perturbation analysis techniques for the nonlinear wave equation are developed for plane waves and then extended to radially symmetric spherical waves and two-dimensional waves radiated by a vibrating plate. The latter leads to demonstration of the phenomenon of self-refraction, in which the rays and wavefronts are modified by the associated pressure and particle velocity fields.

Although the overriding precept of the text is that all topics must be fully explained as they arise, two appendices are provided for further assistance. One is devoted to derivation of the coordinate transformations and vector differential operators in spherical and cylindrical coordinates. The second describes Fourier transforms and their application. Fourier transforms appear in the main body only in the treatment of radiation from an infinite cylinder and from a transducer in a baffle. Those analyses could be addressed by invoking FFT techniques, albeit without the benefit of an algebraic solution. However, Fourier transforms are a ubiquitous thread that runs through the technical literature, and the treatment of this mathematical tool in a sense ties together the dual nature of the time and frequency domains.

The chapters of Volume II of Acoustics-A Textbook for Engineers and Physicists are sequenced in a manner that I believe to be sensible from the viewpoint of ascending difficulty, as well as the technical sophistication. For instance, it would not make sense to study piston radiation in Chap. 8 before the Kirchoff-

Helmholtz integral theorem is developed in Chap. 7. From the opening motivational discussion of waveguides in Chap. 10 to the development of natural cavity modes, there is much reliance on the nature of propagation modes of a waveguide, which is the subject of Chap. 9. Chapter 12 on scattering relies on the KHIT and multipole expansions, which are developed in Chap. 7. Chapters 11 and 13 are exceptions. Both only require concepts developed in the first four chapters. Their placement is based on my perception of what instructors expect to find. My own research in acoustics began in the nonlinear regime, so if I followed my preference, the last chapter might have appeared earlier. In the same vein, I think that exposure of students to ray tracing for heterogeneous media would interest them greatly and thereby serve as a strong motivational tool.

Writing this volume has helped me close gaps in my knowledge of acoustics. I very much enjoyed writing it. If you are a student, I hope you learn much from it and find it useful for your future endeavors. If you are an instructor, I hope that this work captivates you and enhances your teaching efforts. If you are a reader on a self-study path, then I think you will find this volume to be exceptionally helpful and instructive, for I have gone to great effort to provide all you need to pursue your studies.

Jerry H. Ginsberg

## Acknowledgements

I am indebted to many individuals for providing motivation to write these books. Above all, neither would exist if I had not met my good friend, Allan Pierce, when I interviewed in 1980 for a professorship at the Georgia Institute of Technology. Working with him convinced me to extend my knowledge of acoustics beyond the specialized subject of nonlinear acoustics that was part of my early career. Over the years, our discussions were quite revelatory regarding where there were gaps in my knowledge of the subject. Furthermore, I hope he is not offended by this remark, but learning from, and then teaching from, his book convinced me of the necessity that I write Acoustics-A Textbook for Engineers and Physicists.

In the six-year interval, during which I wrote these books what I needed most was assurance that the effort was worth pursuing. Some of my colleagues at Georgia Tech were quite supportive. Karim Sabra convinced me on several occasions that I would be filling an important need. Students solving the homework exercises should thank him because he suggested that the MATLAB code I used should be publically available. Pete Rogers provided extremely useful critical remarks for an early draft of my treatment of geometrical acoustics. My former Ph.D. students, especially J. Gregory McDaniel at Boston University and Kuangchung Wu at the NSWC Carderock Division of the Naval Sea Systems Command, were especially enthusiastic. I also am indebted to those attendees at many Acoustical Society of America meetings who I waylaid to discuss my writing efforts. They are too numerous to list, and I am sure that I have forgotten some names, but I greatly appreciate their attention. I owe Mark Hamilton of the University of Texas at Austin a special debt because he convinced me to participate in the ASA Book program under the aegis of Springer Publishing. Sara Kate Heukerott, my Editor at Springer, was quite understanding of my requests. Her expertise was a great aid as we assembled this project. Some might be surprised at the inclusion of my granddaughter, Leah Morgan Ginsberg, in the list of folks deserving recognition. Early in the writing stage, because she was a proficient clarinetist, I sought her assistance for the discussion of music in Chap. 1. Then at the conclusion, as she approached graduation from Georgia Tech in the G. W. Woodruff School of Mechanical

Engineering, from which I had retired, she served as my sounding board and spokeswoman for students when I deliberated how best to disseminate this work.

In addition to Leah's role, my family was essential to the effort. The forbearance of my wife, Rona, while I focused on writing, ignored other responsibilities, and forgot many things that I still cannot remember, astonishes me, even now that my efforts are over. She went through this experience before when I wrote my prior books on statics, dynamics, and vibrations. However, none of those experiences could have prepared her for the intensity and duration of the present effort. My sons, Mitchell and Daniel, had similar experiences when they lived at home. Although they and their wives, Tracie and Jessica, were not as strongly impacted now, I greatly appreciate their forbearance when I was not as communicative as I should have been. My granddaughters, Leah, Beth, and Abby, inspire me by their dedication to their own activities. I hope that recognition of the pleasure their Papa derived from creating these books will inspire them.

Dunwoody, GA, USA
Jerry H. Ginsberg

## Contents

Volume I: Fundamentals
1 Descriptions of Sound ..... 1
1.1 Harmonic Signals ..... 4
1.1.1 Basic Properties ..... 4
1.1.2 Vectorial Representation ..... 8
1.1.3 Complex Exponential Representation ..... 9
1.1.4 Operations Using Complex Exponentials ..... 12
1.2 Averages ..... 17
1.3 Metrics of Sound ..... 27
1.3.1 Sound Pressure Level ..... 27
1.3.2 Human Factors ..... 32
1.3.3 Frequency Bands ..... 35
1.4 Transfer Between Time and Frequency Domains ..... 41
1.4.1 Fourier Series ..... 43
1.4.2 Discrete Fourier Transforms ..... 52
1.4.3 Nyquist Sampling Criterion ..... 55
1.4.4 Fast Fourier Transforms ..... 59
1.4.5 Evaluation of Time Responses ..... 64
1.5 Spectral Density ..... 72
1.5.1 Definition ..... 72
1.5.2 Noise Models ..... 78
1.6 Closure. ..... 83
1.7 Homework Exercises ..... 83
2 Plane Waves: Time Domain Solutions. ..... 91
2.1 Continuum Equations in One Dimension ..... 92
2.1.1 Conservation of Mass ..... 92
2.1.2 Momentum Equation ..... 94
2.2 Linearization and the One-Dimensional Wave Equation ..... 96
2.3 Equation of State and the Speed of Sound ..... 102
2.4 The d'Alembert Solution ..... 111
2.4.1 Derivation ..... 112
2.4.2 Interpretation. ..... 114
2.4.3 Harmonic Waves ..... 121
2.5 The Method of Wave Images. ..... 123
2.5.1 Initial Value Problem in an Infinite Domain ..... 124
2.5.2 Plane Waves in a Semi-infinite Domain ..... 131
2.5.3 Plane Waves in a Finite Waveguide ..... 145
2.6 Analogous Vibratory Systems ..... 170
2.6.1 Stretched Cable. ..... 170
2.6.2 Extensional Waves in an Elastic Bar ..... 179
2.7 Closure. ..... 183
2.8 Homework Exercises ..... 183
3 Plane Waves: Frequency-Domain Solutions ..... 191
3.1 General Solution ..... 192
3.2 Waveguides with Boundaries ..... 199
3.2.1 Impedance and Reflection Coefficients ..... 200
3.2.2 Evaluation of the Signal ..... 207
3.2.3 Modal Properties and Resonances. ..... 212
3.2.4 Impedance Tubes ..... 223
3.3 Effects of Dissipation ..... 227
3.3.1 Viscosity ..... 229
3.3.2 Thermal Transport ..... 231
3.3.3 Molecular Relaxation ..... 234
3.3.4 Absorption in the Atmosphere and Ocean. ..... 238
3.3.5 Wall Friction. ..... 243
3.4 Acoustical Transmission Lines ..... 248
3.4.1 Junction Conditions ..... 248
3.4.2 Time Domain ..... 252
3.4.3 Frequency-Domain Formulation for Long Segments ..... 257
3.5 Lumped Parameter Models. ..... 272
3.5.1 Approximations for Short Branches ..... 272
3.5.2 Helmholtz Resonator. ..... 278
3.6 Closure. ..... 285
3.7 Homework Exercises ..... 286
4 Principles and Equations for Multidimensional Phenomena ..... 295
4.1 Fundamental Equations for an Ideal Gas ..... 296
4.1.1 Continuity Equation ..... 296
4.1.2 Momentum Equation ..... 298
4.2 Linearization ..... 303
4.3 Plane Waves in Three Dimensions ..... 306
4.3.1 Simple Plane Wave in the Time Domain ..... 307
4.3.2 Trace Velocity ..... 311
4.3.3 Simple Plane Wave in the Frequency Domain ..... 316
4.4 Velocity Potential. ..... 320
4.5 Energy Concepts and Principles ..... 324
4.5.1 Energy and Power ..... 324
4.5.2 Linearization ..... 328
4.5.3 Power Sources ..... 331
4.6 Closure. ..... 341
4.7 Homework Exercises ..... 342
5 Interface Phenomena for Planar Waves ..... 347
5.1 Radiation Due to Surface Waves ..... 347
5.1.1 Basic Analysis ..... 348
5.1.2 Interpretation. ..... 354
5.2 Reflection from a Surface Having a Local Impedance ..... 358
5.2.1 Reflection from a Time-Domain Perspective ..... 358
5.2.2 Reflection from a Frequency-Domain Perspective ..... 365
5.3 Transmission and Reflection at an Interface Between Fluids ..... 373
5.3.1 Time-Domain Analysis ..... 373
5.3.2 Frequency-Domain Analysis ..... 380
5.4 Propagation Through Layered Media ..... 390
5.4.1 Basic Analysis of Three Fluids ..... 390
5.4.2 Multiple Layers. ..... 399
5.5 Solid Barriers ..... 408
5.5.1 General Analysis. ..... 409
5.5.2 Specific Barrier Models ..... 414
5.6 Closure. ..... 424
5.7 Homework Exercises ..... 424
6 Spherical Waves and Point Sources ..... 433
6.1 Spherical Coordinates. ..... 433
6.2 Radially Vibrating Sphere-Time-Domain Analysis ..... 437
6.2.1 General Solution ..... 438
6.2.2 Radiation from a Uniformly Vibrating Sphere ..... 439
6.2.3 Acoustic Field in a Spherical Cavity. ..... 444
6.3 Radially Vibrating Sphere-Frequency-Domain Analysis ..... 453
6.3.1 General Solution ..... 453
6.3.2 Radiation from a Radially Vibrating Sphere ..... 454
6.3.3 Standing Waves in a Spherical Cavity ..... 462
6.4 Point Sources ..... 468
6.4.1 Single Source ..... 468
6.4.2 Green's Function ..... 470
6.4.3 Point Source Arrays ..... 483
6.4.4 Method of Images. ..... 493
6.5 Dipoles, Quadrupoles, and Multipoles ..... 505
6.5.1 The Dipole Field. ..... 506
6.5.2 Radiation from a Translating Rigid Sphere ..... 513
6.5.3 The Quadrupole Field ..... 521
6.5.4 Multipole Expansion ..... 532
6.6 Doppler Effect ..... 538
6.6.1 Introduction ..... 538
6.6.2 Moving Fluid ..... 539
6.6.3 Subsonic Point Source ..... 541
6.6.4 Supersonic Point Source ..... 548
6.7 Closure. ..... 554
6.8 Homework Exercises ..... 554
Index ..... 569
Volume II: Applications
7 Radiation from Vibrating Bodies. ..... 1
7.1 Spherical Harmonics ..... 2
7.1.1 Separation of Variables ..... 2
7.1.2 Description of the Pressure Field ..... 15
7.1.3 Arbitrary Spatial Dependence ..... 16
7.2 Radiation from a Spherical Body ..... 19
7.2.1 Analysis ..... 19
7.2.2 Important Limits ..... 21
7.2.3 Symmetry Plane ..... 29
7.2.4 Interaction with an Elastic Spherical Shell ..... 36
7.3 Radiation from an Infinite Cylinder ..... 53
7.3.1 Separation of Variables ..... 56
7.3.2 Transverse Dependence-Cylindrical Bessel Functions ..... 57
7.3.3 Radiation Due to a Helical Surface Wave ..... 60
7.3.4 Axially Periodic Surface Vibration ..... 68
7.3.5 Finite Length Effects. ..... 81
7.4 Kirchhoff-Helmholtz Integral Theorem ..... 84
7.4.1 Derivation for an Acoustic Cavity ..... 85
7.4.2 Acoustic Radiation into an Exterior Domain. ..... 91
7.5 Numerical Methods for Radiation from Arbitrary Objects ..... 104
7.5.1 Source Superposition ..... 104
7.5.2 Boundary Element Method ..... 108
7.5.3 Finite Element Method ..... 115
7.6 Homework Exercises ..... 125
8 Radiation from a Source in a Baffle ..... 133
8.1 The Rayleigh Integral. ..... 133
8.2 Farfield Directivity ..... 137
8.2.1 Cartesian Coordinate Description ..... 137
8.2.2 Farfield of a Piston Transducer. ..... 144
8.3 Axial Dependence for a Circular Transducer ..... 153
8.4 An Overall Picture of the Pressure Field ..... 158
8.5 Radiation Impedance of a Circular Piston ..... 166
8.6 Time Domain Rayleigh Integral ..... 176
8.7 Homework Exercises ..... 180
9 Modal Analysis of Waveguides ..... 187
9.1 Propagation in a Horn ..... 187
9.1.1 The Webster Horn Equation ..... 188
9.1.2 Exponential Horn ..... 193
9.1.3 Group Velocity ..... 202
9.1.4 WKB Solution for an Arbitrary Horn ..... 207
9.2 Two-Dimensional Waveguides ..... 216
9.2.1 General Solution ..... 217
9.2.2 Rigid Walls ..... 218
9.2.3 Interpretation. ..... 220
9.2.4 Flexible Walls. ..... 224
9.2.5 Orthogonality and Signal Generation ..... 243
9.3 Three-Dimensional Waveguides ..... 253
9.3.1 General Analytical Procedure ..... 253
9.3.2 Rectangular Waveguide ..... 257
9.3.3 Circular Waveguide ..... 267
9.4 Homework Exercises ..... 282
10 Modal Analysis of Enclosures ..... 291
10.1 Fundamental Issues ..... 291
10.1.1 Wall-Induced Signals ..... 291
10.1.2 Source Excitation ..... 293
10.2 Frequency-Domain Analysis Using Forced Cavity Modes ..... 297
10.2.1 Rectangular Enclosures ..... 297
10.2.2 Spherical Cavities ..... 311
10.2.3 Cylindrical Enclosures ..... 323
10.3 Analysis Using Natural Cavity Modes ..... 336
10.3.1 Equations Governing Cavity Modes ..... 337
10.3.2 Orthogonality ..... 338
10.3.3 Analysis of the Pressure Field. ..... 340
10.3.4 Rectangular Cavity ..... 342
10.3.5 Cylindrical Cavity ..... 352
10.3.6 Spherical Cavity ..... 358
10.4 Approximate Methods ..... 363
10.4.1 The Rayleigh Ratio and Its Uses ..... 364
10.4.2 Dowell's Approximation ..... 379
10.5 Homework Exercises ..... 395
11 Geometrical Acoustics ..... 405
11.1 Basic Considerations: Wavefronts and Rays ..... 406
11.1.1 Field Equations for an Inhomogeneous Fluid ..... 409
11.1.2 Reflection and Refraction of Rays ..... 413
11.2 Propagation in a Vertically Stratified Medium ..... 417
11.2.1 Snell's Law for Vertical Heterogeneity ..... 418
11.2.2 Intensity and Focusing Factor. ..... 428
11.3 Arbitrary Heterogeneous Fluids ..... 438
11.3.1 Ray Tracing Equations ..... 438
11.3.2 Amplitude Dependence ..... 443
11.4 Fermat's Principle ..... 463
11.5 Homework Exercises ..... 473
12 Scattering ..... 479
12.1 Background ..... 480
12.2 Scattering by Heterogeneity ..... 482
12.2.1 General Equations ..... 483
12.2.2 The Born Approximation ..... 485
12.3 Rayleigh Scattering Limit. ..... 490
12.3.1 The Rayleigh Limit of the Born Approximation ..... 490
12.3.2 Mismatched Heterogeneous Region ..... 492
12.3.3 Scattering from a Rigid Body ..... 494
12.4 Measurements and Metrics. ..... 505
12.5 High-frequency Approximation ..... 510
12.6 Scattering from Spheres ..... 515
12.6.1 Stationary Spherical Scatterer ..... 515
12.6.2 Scattering by an Elastic Spherical Shell ..... 522
12.7 Homework Exercises ..... 532
13 Nonlinear Acoustic Waves ..... 539
13.1 Riemann's Solution for Plane Waves ..... 540
13.1.1 Analysis ..... 540
13.1.2 Interpretation. ..... 543
13.1.3 Boundary and Initial Conditions ..... 545
13.1.4 Equations of State ..... 549
13.1.5 Quantitative Evaluations ..... 553
13.2 Effects of Nonlinearity ..... 568
13.2.1 Harmonic Generation ..... 568
13.2.2 Shock Formation. ..... 574
13.2.3 Propagation of Weak Shocks ..... 577
13.3 General Analytical Techniques ..... 597
13.3.1 A Nonlinear Wave Equation ..... 598
13.3.2 Frequency-Domain Formulation ..... 602
13.3.3 Regular Perturbation Series Expansion ..... 611
13.3.4 Method of Strained Coordinates ..... 620
13.4 Multidimensional Systems ..... 625
13.4.1 Finite Amplitude Spherical Wave ..... 625
13.4.2 Waves in Cartesian Coordinates ..... 637
13.5 Further Studies ..... 654
13.6 Homework Exercises ..... 658
Appendix A: Curvilinear Coordinates ..... 665
Appendix B: Fourier Transforms ..... 677
Index ..... 689

## About the Author



Jerry H. Ginsberg began his technical education at the Bronx High School of Science, from which he graduated in 1961. This was followed by a B.S.C.E. degree in 1965 from the Cooper Union and an E.Sc.D. degree in engineering mechanics from Columbia University in 1970, where he held Guggenheim and NASA Fellowships. From 1969 to 1973, he was an Assistant Professor in the School of Aeronautics, Astronautics, and Engineering Science at Purdue University. He then transferred to Purdue's School of Mechanical Engineering, where he was promoted to Associate Professor in 1974. In the 1975-1976 academic year, he was a Fulbright-Hayes Advanced Research Fellow at the École Nationale Supérieure d'Électricité et de Mécanique in Nancy, France. He came to Georgia Tech in 1980 as a Professor in the School of Mechanical Engineering, which awarded him the George W. Woodruff Chair in 1989. He retired in June 2008. His prior publications include five textbooks in statics, dynamics, and vibrations, most in several editions, as well as more than one hundred and twenty refereed papers covering these subjects. Dr. Ginsberg became a Fellow of the Acoustical Society of America in 1987 and a Fellow of the American Society of Mechanical Engineers in 1989. The awards and recognitions he has received include Georgia Tech Professor of the Year (1994), ASEE Archie Higdon Distinguished Educator in Mechanics (1998), ASA Trent-Crede Medal (2005),

ASME Per Bruel Gold Medal in Noise Control and Acoustics (2007), and the ASA Rossing Prize in Acoustics Education (2010). In addition to his technical activities, he is an exceptional photographer.

## List of Examples

The name of the MATLAB programs used to solve the respective example are enclosed in parentheses. They are available for download from the Springer website.

| Example 7.1 | (Legendre_function_series.m) Evaluation <br> of a Legendre polynomial series to fit a given function, <br> and evaluation of the series length required to have an |
| :--- | :--- |
| Example 7.2 | error below a specified level. . . . . . . . . . . . . <br> (none) Determination of the spherical harmonic series <br> for the farfield radiated by a combination of a <br> monopole, a dipole, and a longitudinal quadrupole . . . . |


| Example 7.3 | (Cap_in_sphere_radiation.m) Determination <br> of the spherical harmonic series representation of the <br> axisymmetric pressure field radiated by a sphere that <br> has a cap of arbitrary size that executes a translational <br> vibration, and is rigid elsewhere . . . . . . . . . . . . . . . 24 |
| :--- | :--- |

Example 7.4 (Sphere_radiation_baffle_vs_free.m) Evaluation of spherical harmonic series for three cases: a hemisphere that executes a translational vibration in a rigid baffle, the same motion for a pressure-release baffle, and a rigid sphere whose forward half translates while the back half is stationary32

Example 7.5 (Submerged_spher_shell_response.m) Evaluation
of the surface vibration and pressure when a
submerged spherical shell is excited by a harmonic
force concentrated at the polar apex ..... 46
Example 7.6 (Z_rad_helical.m) Determination and evaluation of the radiation impedance for a single helical wave of radial displacement on a cylinder ..... 67

| Example 7.7 | (Line_source_finite.m) Determination of the pressure field radiated by a uniform line source of finite length, and comparison of its field properties to those of an infinite line source |
| :---: | :---: |
| Example 7.8 | (Cello_string_radiation.m) Evaluation of the pressure field radiated by a cello's string based on a model of the string as a cylinder whose cross sections execute translational vibration |
| Example 7.9 | (KHIT_spher_cavity.m) Formulation of the Kirchhoff-Helmholtz integral theorem for a spherical cavity by using the analytical solution to set the pressure and particle velocity at the wall, with special attention to the behavior as the field point approaches the wall. |
| Example 7.10 | (KHIT_cyl_radiation.m) Usage of the radiation impedance of an infinite cylinder as the basis for an approximate formulation of the Kirchhoff-Helmholtz Integral Theorem for a finite cylinder, with emphasis on evaluation of the integral to identify the farfield directivity at several frequencies . |
| Example 8.1 | (Directivity_rectangular_transducer.m) Usage of the Rayleigh integral to evaluate the farfield radiated by a rectangular strip that executes a transverse vibration in a rigid baffle |
| Example 8.2 | (Example_piston_directivities.m) Comparison of the farfield properties for radiation from an oscillating piston or a vibrating membrane in a rigid baffle. |
| Example 8.3 | (Anti_piston_example.m) Determination of the axial dependence of the pressure radiated by a so-called anti-piston, which consists of transducer that has an outer region that vibrates $180^{\circ}$ out-of-phase from the central circle |
| Example 8.4 | (Rayleigh_integral_hemisphere.m) Evaluation of the axial dependence and farfield radiated by a hemisphere in a rigid baffle that translates as a rigid body, based on phase-shifting the normal velocity to place it in the plane of the baffle, followed by comparison of the results to those obtained from the spherical harmonic solution. |
| Example 8.5 | (Piston_square_wave.m) Evaluation of the time dependence of the radiated power and the pressure waveform at a field point on the axis of a circular piston that executes a square wave oscillation. . . . |


| Example 9.1 | Usage of the general solution of the Webster horn equation to evaluate the two-port mobility for an exponential transition between waveguides having different cross-sectional sizes . |
| :---: | :---: |
| Example 9.2 | (Group_velocity_from_data.m) Evaluation of the waveform at two locations in an exponential horn that is driven by a biharmonic input, followed by the usage of those waveforms to determine the group velocity, and comparison of the group velocity to the value derived from the dispersion equation. |
| Example 9.3 | (Webster_horn_numerical_vs_WKB.m) Evaluation of the WKB solution for the pressure distribution in a plane waveguide whose cross-sectional area varies sinusoidally, with a small minimum, followed by comparison of the results to those obtained from a finite difference solution of the Webster horn equation |
| Example 9.4 | (Compliant_2d_waveguide.m) Determination of the phase and group velocity and the transverse mode functions of a two-dimensional waveguide for which one wall is rigid and the other is purely resistive |
| Example 9.5 | (Elastic_2D_waveguide.m) Determination of the eigenvalues of the transverse mode functions and the ratio of plate displacement to the pressure amplitude at selected frequencies for a two-dimensional waveguide whose walls are identical elastic plates . . . . . . . . . . |
| Example 9.6 | (Two_dim_waveguide_excitation.m) Determination of the axial and transverse dependence of pressure amplitude in a two-dimensional waveguide whose floor is rigid and whose upper boundary is a free surface when the excitation stems from a vibrating ribbon transducer aligned horizontally at one end . . |
| Example 9.7 | (Rect_waveguide_line_source.m) Determination of the axial and transverse dependence of pressure amplitude in a three-dimensional waveguide whose sides and floor are rigid and whose upper boundary is a free surface when the excitation stems from a vibrating ribbon transducer aligned vertically. |
| Example 9.8 | (Annular_waveguide_modes.m) Determination of the eigenvalues and transverse mode functions for a waveguide that is the region bounded by concentric rigid cylinders |


| Example 9.9 | (Piston_circular_waveguide.m) Determination <br> of the axial dependence of the pressure amplitude <br> in a circular waveguide when a vibrating piston of <br> lesser diameter is situated concentrically at one end . . . . <br> (Three_D_cavity_force_plate.m) Evaluation | 280 |
| :--- | :--- | :--- |
| Example 10.1 |  |  |
| of the frequency dependence of the force required |  |  |
| to move a massive plate that is part of one wall |  |  |
| in a hard-walled rectangular cavity . . . . . . . . . . . |  |  |
| (Piston_transducer_rectangular_cavity.m) Evaluation |  |  |
| of the frequency dependence of the pressure amplitude |  |  |
| at selected locations and the force required to drive |  |  |$\quad 301$

Example 10.10 (none) Analysis of the modal properties of a system in which an inner sphere is a membrane filled with a fluid, and surrounded by a concentric sphere filled with a different fluid ..... 360
Example 10.11 (none) Usage of the Rayleigh ratio in conjunction with different trial functions to estimate the fundamental natural frequency of a two-dimensional cavity ..... 368
Example 10.12 (Rayleigh_Ritz_Ellipse.m) Evaluation of the natural frequencies of a two-dimensional elliptical cavity according to the Rayleigh-Ritz method, with emphasis on convergence properties and agreement with the separation theorem ..... 374
Example 10.13 (Dowells_approx_2D_cavity.m) Formulation and solution of differential equations derived from Dowell's approximation for a harmonically driven rectangular two-dimensional cavity whose walls have properties that change discontinuously ..... 388
Example 11.1 (Ray_reflection_spherical_cap.m) Evaluation of the rays resulting from incidence of a plane wave on the inner surface of a spherical cap ..... 414
Example 11.2 (Ray_single_stratified_fluid.m) Evaluation of a ray path and propagation time along that ray in a water channel whose sound speed depends parabolicallyon depth424Example 11.3 (Ray_reflections_stratified_fluid.m) Computationof the full set of rays, including those reflected fromthe surface, for a water channel whose sound speeddepends parabolically on depth, and usage of the resultto evaluate a focusing factor.434
Example 11.4 (none) Proof that the ray tracing equations for a channel whose sound speed varies linearly with depth requires that the rays be circles whose center is situated at the extrapolated depth where the sound speed vanishes ..... 441Example 11.5 (Rays_cylindrical_heterog.m) Solution of the raytracing equations for the case of a plane wave thatpasses transversely through a cylindrically symmetricheterogeneous region, with special attention to thecaustic that is created, followed by applicationof the transport equation to determine the variationof sound along a wavefront451
Example 11.6 (none) Usage of geometrical acoustics to evaluate the eigenrays for a point source below the free surface in a homogeneous channel, and comparison of the result to that obtained from the method of images. ..... 459
Example 11.7 (none) Application of Fermat's principle to determine the eigenrays for propagation from a point source in water channel of constant depth bounded above by air in a half-space, and below by a liquid sediment in a half-space ..... 469
Example 12.1 (Born_approximation_cylinder.m) Evaluation of the angular dependence of the pressure scattered by a cylindrical region of heterogeneity according to the Born approximation ..... 486
Example 12.2 (Scattering_rigid_body_Rayleigh.m) Application of the Rayleigh approximation of scattering to evaluate bistatic scattering of a rigid body ..... 497
Example 12.3 (Scattering_movable_rigid_disk.m) Evaluation of low-frequency scattering at arbitrary incidence on a rigid circular disk that is free to move ..... 502
Example 12.4 (Submerged_spher_shell_scattering.m) Evaluation of the backscatter cross section and total scattering cross section for a spherical shell in water and in the atmosphere ..... 527
Example 13.1 (none) Application of a graphical construction method to evaluate the nonlinear propagation of two plane wave pulses in the case where the initial waveform of the second is inverted from that of the first ..... 556
Example 13.2 (Nonlin_harmonic_plane_wave.m) Computation of the nonlinear propagation of the nonlinear plane wave generated by a transient sinusoidal excitation using a simple algorithm ..... 560
Example 13.3 (Nonlinear_wave_Newtons_method.m) Application of Newton's method to evaluate the Riemann solution for a nonlinear plane wave when the excitation is the sum of a fundamental and second harmonic, followed by FFT analysis to determine the frequency spectrum ..... 565
Example 13.4 (Nonlin_harmonic_wave_shock_fitting.m) Implementation of the equal-area rule to fit shocks into the Riemann solution for an initially sinusoidal nonlinear plane wave ..... 587
Example 13.5 (Frequency_domain_nonlin_ODEs.m) Solution of the coupled differential equations governing the nonlinear propagation of the Fourier series coefficients of an initially sinusoidal plane wave in the case where the attenuation coefficient increases as the square of frequency ..... 607
Example 13.6 (none) Formulation of a regular perturbation series to analyze the nonlinear wave generated by a periodic sequence of beats in order to identify the growth of harmonic and intermodulation distortion ..... 607
Example 13.7 (Nonlin_spher_wave_2_harmonics.m) Evaluation of the waveform of a nonlinear spherical wave generated by boundary vibration that is the sum of a fundamental and a third harmonic, and comparison of the behavior of a nonlinear plane wave generated by the same excitation ..... 633
Example 13.8 (Nonlin_2D_waveforms.m) Application of Newton's method to solve the transcendental equations resulting from singular perturbation analysis of the two-dimensional nonlinear plane wave radiated by a periodically supported plate ..... 651

## Chapter 7 <br> Radiation from Vibrating Bodies

The previous chapters were devoted to the analysis and interpretation of specific types of acoustic waves. In this chapter, the central theme is phenomenological. We will consider a variety of situations in which the task is to determine the acoustic field generated by a body that executes a specified vibration. These are the problems in acoustic radiation.

This chapter opens with an investigation of spherical waves that are not radially symmetric. A description of waves in terms of spherical harmonics, which may be considered to be the extension of Fourier series to a spherical geometry, will be applied to treat radiation from a sphere whose surface velocity pattern is arbitrary. At the same time, it is reasonable to ask how a velocity pattern can be imposed in reality. An object moves because forces are applied to it. The analysis of the vibration of a body excited by forces is a fundamental subject unto itself. In an acoustic system, part of the force system acting on a body is the surface pressure. It follows that to study the vibration problem, we must determine the radiated field, and that determination of the radiated field requires that we solve the vibration problem. In other words, the two problems are coupled. The joint formulation and solution of this coupled problem are the realm of structural acoustics. The behavior of a spherical shell immersed in a fluid will be examined in order to shed light on the analytical approach and phenomena associated with this subject.

To explore the field radiated by an infinitely long vibrating cylinder, we will develop the general solution of the Helmholtz equation in cylindrical coordinates. Spheres and infinitely long cylinders are just about the only geometrical configurations that are amenable to analysis. Their solutions shed much light on fundamental phenomena, but numerical simulation tools are required to address realistic systems. One such formulation is boundary elements, which is based on a fundamental concept called the Kirchhoff-Helmholtz integral theorem. This theorem, which we shall abbreviate to KHIT, gives the pressure at a field point in terms of an integral of the pressure and velocity on the surface of the radiating body. After the theorem is derived, its implementation as a boundary element formulation will be described. Two other numerical techniques for analyzing acoustic radiation are source superposition, which uses a set of sources to replace the vibrating body, and finite elements,
which requires adaptation to fit the nature of acoustic radiation. These developments close this chapter.

### 7.1 Spherical Harmonics

When the surface of a sphere vibrates in an irregular pattern, the radial velocity is an arbitrary function of the polar and azimuthal angles. Given the difficulties that arose for a radially symmetric wave in the time domain, it should not be surprising that we will only pursue a frequency domain analysis. Furthermore, in order to avoid excessive entanglement in the techniques of mathematical analysis, we will restrict our attention to axisymmetric situations. The consequence is that the pressure is $p=\operatorname{Re}[P(r, \psi) \exp (i \omega t)]$. Although this will limit the generality of the situations we can address, it will not restrict the nature of the physical phenomena that arise.

### 7.1.1 Separation of Variables

Our objective is to determine the complex pressure amplitude $P(r, \psi)$ that is the general solution of the Helmholtz equation. The analysis will use the method of separation of variables. It begins with an ansatz that represents the dependent variable as a product of functions of each of the independent variables. In other words,

$$
\begin{equation*}
P=\mathcal{F}(r) \mathcal{G}(\psi) \tag{7.1.1}
\end{equation*}
$$

It is helpful to rewrite the Helmholtz equation in axisymmetric spherical coordinates, Eq. (A.1.16), as

$$
\begin{equation*}
\frac{1}{r} \frac{\partial^{2}}{\partial r^{2}}(r P)+\frac{1}{r^{2} \sin \psi} \frac{\partial}{\partial \psi}\left(\sin \psi \frac{\partial P}{\partial \psi}\right)+k^{2} P=0 \tag{7.1.2}
\end{equation*}
$$

When we substitute the factorized representation of $P$ into this equation, and divide by $\mathcal{F G} / r^{2}$, the result is an equation in which each side is a function of a different position variable, specifically,

$$
\begin{equation*}
\frac{r}{\mathcal{F}} \frac{d^{2}}{d r^{2}}(r \mathcal{F})+k^{2} r^{2}=-\frac{1}{\mathcal{G} \sin \psi} \frac{d}{d \psi}\left(\sin \psi \frac{d \mathcal{G}}{d \psi}\right) \tag{7.1.3}
\end{equation*}
$$

The argument now is that the only way a function of $r$ can equal a function of $\psi$ is if they both equal some constant, which we denote as $\mathcal{C}$. Thus, we obtain two ordinary differential equations,

$$
\begin{gather*}
\frac{d^{2} \mathcal{F}}{d r^{2}}+\frac{2}{r} \frac{d \mathcal{F}}{d r}+\left(k^{2}-\frac{C}{r^{2}}\right) \mathcal{F}=0  \tag{7.1.4}\\
\frac{1}{\sin \psi} \frac{d}{d \psi}\left(\sin \psi \frac{d \mathcal{G}}{d \psi}\right)+\mathcal{C G}=0
\end{gather*}
$$

These equations are solved independently. We begin with $\mathcal{G}(\psi)$, because doing so will establish the separation constant $\mathcal{C}$.

Actually, we will find that there are an infinite number of $\mathcal{C}$ values. The set of corresponding $\mathcal{G}$ functions are referred to as spherical harmonics. The reason for this terminology stems in part from the oscillatory nature of the functions, and also from the fact that they are the basis functions for a Fourier-like series describing an arbitrary function of the spherical angles. Because our interest is in axisymmetric situations, the spherical harmonics derived here are a subset of the general set of functions, which depend on $\theta$, as well as $\psi$.

## Dependence on the Polar Angle

The presence of sinusoidal terms in the coefficients of the second of Eq.(7.1.4) complicates the task of finding a solution. Fortunately, a transformation from $\psi$ to a nondimensional variable $\eta$ converts the differential equation to a form that has been widely studied. The transformation is

$$
\begin{equation*}
\eta=\cos \psi \Longrightarrow \sin \psi=\left(1-\eta^{2}\right)^{1 / 2} \tag{7.1.5}
\end{equation*}
$$

The derivatives become

$$
\begin{align*}
\frac{d \mathcal{G}}{d \psi} & =-\sin \psi \frac{d \mathcal{G}}{d \eta} \\
\frac{1}{\sin \psi} \frac{d}{d \psi}\left(\sin \psi \frac{d \mathcal{G}}{d \psi}\right) & =(\sin \psi)^{2} \frac{d^{2} \mathcal{G}}{d \eta^{2}}-2 \cos \psi \frac{d \mathcal{G}}{d \eta} \tag{7.1.6}
\end{align*}
$$

We use the transformation to eliminate $\psi$ from the separated equation, which leads to Legendre's equation,

$$
\begin{equation*}
\left(1-\eta^{2}\right) \frac{d^{2} \mathcal{G}}{d \eta^{2}}-2 \eta \frac{d \mathcal{G}}{d \eta}+C \mathcal{G}=0 \tag{7.1.7}
\end{equation*}
$$

One of the most important aspects of this equation is that the coefficient of its highest derivative vanishes. Differential equations having this property contain a solution that is singular. Dealing with this possibility will lead to an expression for $\mathcal{C}$. The fact that the coefficients in this equation are powers of $\eta$ suggests that $\mathcal{G}$ may be represented as a power series. Thus, we try

$$
\begin{equation*}
\mathcal{G}=\sum_{j=0}^{\infty} a_{j} \eta^{j} \tag{7.1.8}
\end{equation*}
$$

The next step is to substitute this form into Legendre's equation, and require that it be satisfied independently for each power of $\eta$. The details of these operations may be found in most textbooks on ordinary differential equations. The result is that the coefficients are obtained as a recurrence relation that gives a higher order coefficient in terms of a lower order value. The specific relation is

$$
\begin{equation*}
a_{j+2}=\frac{j(j+1)-\mathcal{C}}{(j+2)(j+1)} a_{j}, \quad j=0,1,2, \ldots \tag{7.1.9}
\end{equation*}
$$

The implication of this relation is that $a_{0}$ and $a_{1}$ are arbitrary and independent, from which successive coefficients are found as factors multiplying the $a_{0}$ and $a_{1}$ values.

If $C$ is arbitrarily selected, it can be proven that the infinite series for $\mathcal{G}$ does not converge at $n= \pm 1$. This is how the singular nature of Eq. (7.1.7) is manifested. The locations at which $|\eta|=1$ are the poles of the spherical coordinate system, but there is no physical aspect of the poles that would cause a singular response. The only way to avoid this behavior and thereby obtain a solution that is valid everywhere, is to select $\mathcal{C}$ such that the series becomes a finite sum. This happens if $\mathcal{C}$ makes one of the $a_{j}$ coefficients to vanish, because the recurrence relation will lead to the higher coefficients being zero.

If we wish that the $m$ th coefficient be the last one that is nonzero, then we would select

$$
\begin{equation*}
\mathcal{C}_{m}=m(m+1) \tag{7.1.10}
\end{equation*}
$$

Each value of $m$ leads to a different solution of Legendre's equation because $m$ is the degree of the polynomial. The solutions are written as $P_{m}(\eta)$, which denotes that they are Legendre polynomials of degree $m$. Their definition is

$$
\begin{equation*}
P_{m}(\eta)=\sum_{j=0}^{\text {floor }(m / 2)} \frac{(-1)^{j}(2 m-2 j)!}{2^{m} j!(m-j)!(m-2 j)!} \eta^{m}-2 j \tag{7.1.11}
\end{equation*}
$$

One should exercise caution whether this formula is used for computations when $m$ is large, because the factorial terms might be larger than the largest computable floating point number for the particular computer architecture. The lowest degree polynomials are

$$
\begin{gather*}
P_{0}(\eta)=1, \quad P_{1}(\eta)=\eta, \quad P_{2}(\eta)=\frac{1}{2}\left(3 \eta^{2}-1\right)  \tag{7.1.12}\\
P_{3}(\eta)=\frac{1}{2}\left(5 \eta^{3}-3 \eta\right)
\end{gather*}
$$

The substitution of $\eta=\cos \psi$ leads to $P_{m}(\cos \psi)$, which is a Legendre function of $\psi$. The first few functions are

$$
\begin{align*}
P_{0}(\cos \psi)=1, & P_{1}(\cos \psi)=\cos \psi, \quad P_{2}(\cos \psi)=\frac{1}{4}(3 \cos 2 \psi+1)  \tag{7.1.13}\\
& P_{3}(\cos \psi)=\frac{1}{8}(5 \cos 3 \psi+\cos \psi)
\end{align*}
$$

To obtain these forms directly from the Legendre polynomial with $\eta=\cos \psi$, one must apply various trigonometric identities for powers of $\cos \psi$. By virtue of their dependence on $\cos \psi$, the Legendre functions are periodic in $\Delta \psi=2 \pi$. Consequently, $P_{m}(\cos \psi)$ may be represented by a Fourier series. The handbook edited by Abramowitz and Stegun ${ }^{1}$ is an excellent resource for the properties of Legendre functions.

Legendre polynomials are plotted in the upper set of graphs in Fig.7.1. The even degrees are symmetric functions with respect to the origin, while odd degrees correspond to antisymmetric functions. As $m$ increases, $P_{m}(\eta)$ oscillates over a smaller scale, but these oscillations are not harmonic. Recall that the coefficients $a_{0}$ and $a_{1}$ in Eq. (7.1.8) may be selected arbitrarily. The plotted functions have been defined such that $\left|P_{m}( \pm 1)\right|=1$. The lower set of graphs in Fig. 7.1 shows that replacing $\eta$ with $\cos (\psi)$ to form a Legendre function does not fundamentally alter the appearance of the function at any degree. However, it does have the effect of compressing the region near the midpoint, which now is $\psi=\pi / 2$, and expanding the region near the limits at $\eta=0$ and $\pi$. The transformation from $\eta$ to $\psi$ leads to $(d / d \psi) P_{m}(\cos (\psi))$ being identically zero at $\psi=0$ and $\pi$. It follows that setting $\mathcal{G}$ to a Legendre function in the separation of variables form, Eq.(7.1.1), will lead to $d P / d \psi$ being zero identically


Fig. 7.1 Low-degree Legendre polynomials as a function of $\eta$, and Legendre functions as a function of $\psi$

[^0]at the poles. The continuity equation requires that an axisymmetric field meets this condition.

The Legendre polynomials have some useful properties. The first are a set of recurrence relations for functions and derivatives at different degrees,

$$
\begin{gather*}
(m+1) P_{m+1}(\eta)=(2 m+1) \eta P_{m}(\eta)-m P_{m-1}(\eta) \\
\eta \frac{d}{d \eta} P_{m}(\eta)-\frac{d}{d \eta} P_{m-1}(\eta)=m P_{m}(\eta) \\
\frac{d}{d \eta} P_{m+1}(\eta)-\frac{d}{d \eta} P_{m-1}(\eta)=(2 m+1) P_{m}(\eta)  \tag{7.1.14}\\
\left(1-\eta^{2}\right) \frac{d}{d \eta} P_{m}(\eta)=-m \eta P_{m}(\eta)+m P_{m-1}(\eta)
\end{gather*}
$$

A different definition of the polynomials is offered by Rodrigue's formula,

$$
\begin{equation*}
P_{m}(\eta)=\frac{1}{2^{m} m!} \frac{d^{m}}{d \eta^{m}}\left[\left(\eta^{2}-1\right)^{m}\right] \tag{7.1.15}
\end{equation*}
$$

An important property of Legendre polynomials is that they are mutually orthogonal. The polynomials are defined over the range $-1 \leq \eta \leq 1$, so orthogonality refers to an inner product defined as an integral over that interval. The orthogonality property states that if $m$ and $j$ are any two indices, then

$$
\begin{equation*}
\int_{-1}^{1} P_{m}(\eta) P_{j}(\eta) d \eta=\int_{0}^{\pi} P_{m}(\cos \psi) P_{j}(\cos \psi) \sin \psi d \psi=\frac{2}{2 m+1} \delta_{j n} \tag{7.1.16}
\end{equation*}
$$

where $\delta_{j m}$ is the Kronecker delta, which equals one if $j=m$ and zero otherwise.
Recall that the orthogonality of harmonic functions is one of the primary reasons that a Fourier series has great utility. Orthogonality has a similar usefulness when a variable is represented by a Legendre series. In a Fourier series, the harmonic functions are said to be the basis functions and the series coefficients are "distances," in a linear algebra sense, in the direction of each harmonic function. A Legendre series is defined similarly, with $P_{m}(\eta)$ as the basis functions, specifically,

$$
\begin{equation*}
F(\eta)=\sum_{m=0}^{\infty} F_{m} P_{m}(\eta) \text { or } F(\psi)=\sum_{m=0}^{\infty} F_{m} P_{m}(\cos \psi) \tag{7.1.17}
\end{equation*}
$$

The similarity of this representation to a Fourier series causes some individuals to say that the preceding is a generalized Fourier series.

Suppose we wish to determine the $F_{m}$ coefficients associated with a specified $F$ function. For an acoustics problem, $\psi$ is the independent variable, so we shall work
with the second form in Eq. (7.1.17). Both sides of the equation are multiplied by $P_{j}(\cos \psi) \sin \psi$, where $j$ is a designated index. Then, both sides are integrated over $0<\psi<\pi$. We bring the integration inside the summation, with the result that

$$
\begin{equation*}
\int_{0}^{\pi} F(\psi) P_{j}(\cos \psi) \sin \psi d \psi=\sum_{m=0}^{\infty} F_{m} \int_{0}^{\pi} P_{m}(\cos \psi) P_{j}(\cos \psi) \sin \psi d \psi \tag{7.1.18}
\end{equation*}
$$

The orthogonality property states that all integrals vanish, except for the one where the summation index $m$ equals the selected value of $j$. In that case, the integral is $2 /(2 j+1)$. Thus, the orthogonality property filters from the sum all except the $j$ th term. The result is that

$$
\begin{equation*}
F_{j}=\frac{2 j+1}{2} \int_{0}^{\pi} F(\psi) P_{j}(\cos \psi) \sin \psi d \psi \tag{7.1.19}
\end{equation*}
$$

One may attempt to evaluate the integral analytically, possibly by replacing $P_{j}(\cos \psi)$ with its definition as a polynomial, or else with the aid of a table of integrals. An alternative is to use numerical methods to evaluate the integral.

EXAMPLE 7.1 Legendre polynomials have been used to describe crystal vibrations. ${ }^{2}$ That analysis used a series of Legendre polynomials to represent the dependence of displacement $u$ on distance $y$ from the midplane of a rectangular plate. The graph depicts $u=0.2 \tanh ^{-1}(1.98 y / h)$, where $h$ is the plate's thickness. Determine the Legendre series coefficients that represent this function. Identify at what degree the series may be truncated such that its evaluation will differ everywhere from the actual by less than $2 \%$ of the maximum value.


[^1]
## Significance

In addition to showing how we may determine a Legendre series numerically, this example will give some insight to the convergence properties of this type of series.

## Solution

To use Eq. (7.1.19), the range of integration must be $-1 \leq \eta \leq 1$, but the range for the plate is $-h / 2 \leq y \leq h / 2$. Thus, we set $2 y / h=\eta$, so that $u(\eta)=0.2 \tanh ^{-1}(0.99 \eta)$. Because $\eta$ is the dependent variable, we use the first form in Eq. (7.1.16). This entails multiplying the series description of $u$ by $P_{j}(\eta)$, then integrating over $-1<\eta<1$. The result is

$$
F_{j}=\left(\frac{2 j+1}{2}\right) 0.2 \int_{-1}^{1} \tanh ^{-1}(0.99 \eta) P_{j}(\eta) d \eta
$$

There is no simple way to evaluate this integral analytically, so we turn to numerical methods. The details depend on the software to be used, but the basic elements are the same. A routine to evaluate the Legendre functions at a set of values of $\eta$ is required, as well as a routine that performs numerical integration. If one wishes to write their own procedures, then Eq.(7.1.11) may be coded, because the degree of the Legendre polynomial will not be extraordinarily high. Numerical integration may be carried out by Simpson's rule, provided one verifies that the sampling interval is sufficiently small. The present results were obtained with MATLAB, which contains a function routine legendre.m that evaluates $P_{m}(x)$ for a vector of $x$ values. However, it cannot be used directly because it also evaluates associated Legendre functions. They are described in Sect. 7.1.3 as functions that would arise if we were to investigate spherical waves that are not axisymmetric. The nature of legendre.m is that it returns an array having $m+1$ rows and $n$ columns, where $m$ is the degree number and $n$ is the number of elements of $x$. The first row of this array holds the $P_{m}(x)$ values, while the other rows hold the corresponding associated functions. Thus, we evaluate the Legendre polynomials by creating a MATLAB function leg.m that retains only the first row, as follows:

```
function P_m = leg(x, m); full_set = legendre(m, x); P_m = full_set(1,:)
```

The integral was evaluated with one of MATLAB's numerical integration routines, quadl.m. Implementation of this scheme requires that we define another function to evaluate the integrand. It is a good idea to verify that everything has been implemented correctly by following the same procedure to evaluate the integral in the orthogonality condition.

The first few coefficients obtained from the numerical integration are $F_{0}=0$, $F_{1}=0.2869, F_{2}=0, F_{3}=0.0972, F_{4}=0$. All coefficients having even subscripts are zero because $u(\eta)$ is an odd function with respect to $\eta=0$, and that is the symmetry of the odd-degree Legendre polynomials. To examine the convergence of the series, the odd-subscripted $F_{m}$ are plotted in Fig. 2. It can be seen that the coefficients decrease rapidly as $m$ increases.


Figure 2.

It is requested that truncation of the Legendre series not lead to an error greater than $2 \%$ of the maximum value of $u$, which is 0.529 at $y=0.5$. This requires that the series be synthesized at the full range of $y$. An estimate for the series length is obtained from the observation that the maximum magnitude of $P_{m}(x)$ is one, so inclusion of a term $F_{N} P_{m}(x)$ in the Legendre series can change the series sum by no more that $\left|F_{n}\right|$. It is specified that the error should be no more that $2 \%$ of 0.529 , so we search for the smallest Legendre degree $N$ at which $\left|F_{N}\right|<0.0106$. This leads to $N=13$. However, a certain error level in the coefficients might not be manifested by the same error when the series is synthesized.

In general, a efficient way in which a series may be evaluated at many points is to write it as a matrix product. Toward that end, we let $\{y\}$ denote a vector of $y$ values, so that $\{\eta\}=(2 / h)\{y\}$. Then, we form a rectangular array $[L]$ whose element at row $j$ and column $m$ is $P_{m}\left(\eta_{j}\right)$. The product of row $j$ of $[L]$ and a vector $\{F\}$ is the same as the series evaluation of $u\left(y_{j}\right)$. Thus, the full set of $\{u\}$ values is found by evaluating $[L]\{F\}$. Figure 3 displays the error, defined as $\left|u\left(y_{j}\right)-u_{\text {series }}\left(y_{j}\right)\right| / u(y=h / 2)$ for $N=13$. The convergence criterion is met away from $y= \pm h / 2$, but it is $4 \%$ at the limits. A larger $N$ is required to satisfy it everywhere. Successive evaluations at increasing $N$ revealed that $N=19$ is the smallest degree at which truncation of the series will not lead to an error exceeding $2 \%$ at any $y$.


Figure 3.

## Radial Dependence

The separation of variables constant $C$ depends on the degree $m$ of the Legendre polynomial, as given by Eq. (7.1.10). Thus, each $m$ leads to a different radial function.

Each is governed by the first of Eq.(7.1.4), which now is

$$
\begin{equation*}
\frac{d^{2} \mathcal{F}}{d r^{2}}+\frac{2}{r} \frac{d \mathcal{F}}{d r}+\left(k^{2}-\frac{m(m+1)}{r^{2}}\right) \mathcal{F}=0 \tag{7.1.20}
\end{equation*}
$$

At this juncture, there are a number of ways by which $\mathcal{F}$ may be determined. One approach is to implement the method of Frobenius, which is a standard tool for solving ordinary differential equations whose coefficients are not constant. The ansatz for this method is a power series multiplied by a factor that accounts for singular behavior. Another approach is suggested by the fact that far from a radiating body, spherical spreading requires that the solution decays reciprocally with $r$. We can anticipate this effect by changing the dependent variable according to

$$
\begin{equation*}
\mathcal{F}=\Phi(r) / r \tag{7.1.21}
\end{equation*}
$$

The differential equation becomes

$$
\begin{equation*}
\frac{d^{2} \Phi}{d r^{2}}+\left[k^{2}-\frac{m(m+1)}{r^{2}}\right] \Phi=0 \tag{7.1.22}
\end{equation*}
$$

This is promising because it tells us that $\Phi$ tends to a harmonic function at larger $r$. Nevertheless, this equation does not appear to be easier to solve than the equation for $F$. Our knowledge of acoustics principles suggests how to proceed. When $m=0$, Eq.(7.1.20) reduces to the Helmholtz equation in the case of a radially symmetric field. Thus, for $m=0$, we know that $\mathcal{F}=B e^{ \pm i k r} / r$. Furthermore, regardless of $m$, we know that the farfield pressure is a spherically spreading wave with an angular directivity. The observations suggest that we change the dependent variable such that

$$
\begin{equation*}
\mathcal{F}=\frac{1}{r} F_{m}(r) e^{-i k r} \tag{7.1.23}
\end{equation*}
$$

Determination of the function $F_{m}(r)$ is expedited by writing the differential equation as

$$
\begin{equation*}
\frac{1}{r}\left[\frac{d^{2}}{d r^{2}}(r \mathcal{F})+\left(k^{2}-\frac{m(m+1)}{r^{2}}\right)(r \mathcal{F})\right]=0 \tag{7.1.24}
\end{equation*}
$$

When the representation of $\mathcal{F}$ is substituted into this equation, the complex exponential cancels. What remains is an ordinary differential equation for $F_{m}$,

$$
\begin{equation*}
\frac{d^{2} F_{m}}{d(k r)^{2}}-2 i \frac{d F_{m}}{d(k r)}-\frac{m(m+1)}{(k r)^{2}} F_{m}=0 \tag{7.1.25}
\end{equation*}
$$

This might not seem to be any more conducive to an analytical solution than the original equation governing $\mathcal{F}$, but it is. In the farfield, $F_{m}$ must approach a constant value, which suggests a series representation in powers of $1 /(k r)$, that is,

$$
\begin{equation*}
F_{m}=\sum_{j=0}^{\infty} \frac{a_{j}}{(k r)^{j}} \tag{7.1.26}
\end{equation*}
$$

The coefficients $a_{j}$ are determined by substituting the series into Eq.(7.1.25), followed by equating the collected coefficients of each power of $1 /(k r)$ to zero. These operations lead to a recurrence relation for $a_{j+1}$ as a proportionality to $a_{j}$, starting with $j=0$. Examination of this relation shows that it gives $a_{j}=0$ if $j>m$. In other words, although we allowed for an infinite length series, the number of terms must be one greater than the index $m$ for the separation constant $\mathcal{C}$. The solution for $\mathcal{F}$ corresponding to each $F_{m}$ is the second kind of spherical Hankel function of order $m$. A detailed analysis would lead to

$$
\begin{equation*}
\mathcal{F}=h_{m}^{(2)}(k r)=i^{m+1} \frac{e^{-i k r}}{k r} \sum_{j=0}^{m} \frac{(m+j)!}{j!(m-j)!}\left(\frac{1}{2 i k r}\right)^{j} \tag{7.1.27}
\end{equation*}
$$

In this form, the starting coefficient $a_{0}$ has been set to one. The lowest order functions are

$$
\begin{gather*}
h_{0}^{(2)}(k r)=\frac{i e^{-i k r}}{k r}, \quad h_{1}^{(2)}(k r)=-\frac{e^{-i k r}}{k r}\left(1-\frac{i}{k r}\right)  \tag{7.1.28}\\
h_{2}^{(2)}(k r)=-\frac{i e^{-i k r}}{k r}\left(1-\frac{3 i}{k r}-\frac{3}{(k r)^{2}}\right)
\end{gather*}
$$

At this juncture, the logical question is: Why is this function called the "second kind"? We found it first, so why is it not the "first kind"? The answer lies in the fact that we have used $e^{i \omega t}$ to represent a time function. The function associated with the negative exponential temporal representation by convention is denoted as $h_{m}^{(1)}(k r)$, which is the first kind of spherical Hankel function of order $m$. Because the proper representation of a harmonic function adds $e^{-i \omega t}$ and $e^{i \omega t}$ terms, the coefficients of the exponential must be complex conjugates in order that the time dependence be real. Hence, it must be that

$$
\begin{equation*}
h_{m}^{(1)}(k r)=h_{m}^{(2)}(k r)^{*} \tag{7.1.29}
\end{equation*}
$$

This definition of another solution could have been anticipated from the fact that the differential equation whose solution we sought, Eq. (7.1.20), is real. If a complex function satisfies a linear differential equation whose coefficients are real, then its complex conjugate must do so as well because the real and imaginary parts of the differential equation must be satisfied independently. The same reasoning leads to the recognition that the real and imaginary parts of the spherical Hankel functions must satisfy the equation independently. The real part is the spherical Bessel function of
order $m$, denoted $j_{m}(k r)$, and the imaginary part is the spherical Neumann function of order $m$, denoted $n_{m}(k r)$,

$$
\begin{equation*}
j_{m}(k r)=\operatorname{Re}\left(h_{m}^{(2)}(k r)\right), \quad n_{m}(k r)=-\operatorname{Im}\left(h_{m}^{(2)}(k r)\right) \tag{7.1.30}
\end{equation*}
$$

A different way of viewing these definitions is that

$$
\begin{gather*}
h_{m}^{(2)}(k r)=j_{m}(k r)-i n_{m}(k r), \quad h_{m}^{(1)}(k r)=j_{m}(k r)+i n_{m}(k r) \\
j_{m}(k r)=\frac{1}{2}\left[h_{m}^{(1)}(k r)+h_{m}^{(2)}(k r)\right], \quad n_{m}(k r)=\frac{1}{2 i}\left[h_{m}^{(1)}(k r)-h_{m}^{(2)}(k r)\right] \tag{7.1.31}
\end{gather*}
$$

The lowest order functions are

$$
\begin{array}{rlrl}
j_{0}(k r)= & \frac{\sin (k r)}{k r} & n_{0}(k r)=-\frac{\cos (k r)}{k r} \\
j_{1}(k r)= & \frac{\sin (k r)}{(k r)^{2}}-\frac{\cos (k r)}{k r} & n_{1}(k r)=-\frac{\cos (k r)}{(k r)^{2}}-\frac{\sin (k r)}{k r} \\
j_{2}(k r)= & \left(\frac{3}{(k r)^{3}}-\frac{1}{k r}\right) \sin (k r) & n_{2}(k r)= & -\left(\frac{3}{(k r)^{3}}-\frac{1}{k r}\right) \cos (k r) \\
& -\frac{3}{(k r)^{2}} \cos (k r) & & -\frac{3}{(k r)^{2}} \sin (k r)
\end{array}
$$

When we wish to refer to these solutions generically, we will say that they are spherical Bessel functions.

Most mathematical software provides routines to evaluate Bessel functions $J_{\nu}(x)$ and Neumann functions $N_{\nu}(x) .{ }^{3}$ Bessel's equation arises in problems featuring a planar circular geometry. For this reason, the functions $J_{\nu}(x)$ and $N_{\nu}(x)$ are said to be a set of cylindrical Bessel functions. The cylindrical and spherical functions are related by

$$
\begin{equation*}
j_{m}(x)=\left(\frac{\pi}{2 x}\right)^{1 / 2} J_{m+1 / 2}(x), \quad n_{m}(x)=\left(\frac{\pi}{2 x}\right)^{1 / 2} N_{m+1 / 2}(x) \tag{7.1.33}
\end{equation*}
$$

In the event that the available software only evaluates cylindrical Bessel function, we would evaluate the cylindrical functions for order $m+1 / 2$ and argument $k r$ and then multiply the result by $\pi^{1 / 2} /(2 k r)^{1 / 2}$. The spherical Hankel functions may be found from Eq. (7.1.31).

The low-order spherical Bessel and Neumann functions are graphed in Fig. 7.2. The abscissa for the Neumann functions is split because that function is singular at

[^2]

Fig. 7.2 Dependence of the spherical Bessel and Neumann functions on the value of its order $m$ and argument $x$
the origin. The behavior for small $x$ of $n_{m}(x)$ is quite different from that of $j_{m}(x)$, but both become oscillatory with decreasing amplitude as $x$ increases.

These observations are borne out by the asymptotic approximations for small and large arguments. These behaviors for all functions may be extracted by the analysis of Eq. (7.1.27), combined with application of Eqs.(7.1.29) and (7.1.30). When the argument is very small, $x \ll 1$, the results are

$$
\left.\begin{array}{c}
h_{m}^{(2)}(x)=h_{m}^{(1)}(x)^{*} \rightarrow i \frac{(2 m)!}{m!2^{m}}\left(\frac{1}{x}\right)^{m+1}  \tag{7.1.34}\\
j_{m}(x) \rightarrow \frac{m!(2 x)^{m}}{(2 m+1)!} \\
n_{m}(x) \rightarrow-\frac{(2 m)!}{m!2^{m}}\left(\frac{1}{x}\right)^{m+1}
\end{array}\right\} \text { as } x \rightarrow 0
$$

The trends for large arguments are

$$
\left.\begin{array}{rl}
h_{m}^{(2)}(x) & =h_{m}^{(1)}(x)^{*} \rightarrow \frac{1}{x} e^{-i(x-(m+1) \pi / 2)}  \tag{7.1.35}\\
j_{m}(x) & \rightarrow \frac{1}{x} \cos \left(x-\frac{m+1}{2} \pi\right) \\
n_{m}(x) & \rightarrow \frac{1}{x} \sin \left(x-\frac{m+1}{2} \pi\right)
\end{array}\right\} \text { as } x \rightarrow \infty
$$

An important aspect of the preceding is that they require that the argument $x$ be larger than the order $m$. Another asymptotic representation covers the situation where the order is large,

$$
\left.\begin{array}{c}
h_{m}^{(2)}(x)=h_{m}^{(1)}(x)^{*} \rightarrow i\left(\frac{2}{(2 m+1) x}\right)^{1 / 2}\left(\frac{2 m+1}{e x}\right)^{(m+1 / 2)} \\
j_{m}(x) \rightarrow\left(\frac{\pi}{2 x}\right)^{1 / 2} \frac{1}{(\pi(2 m+1))^{1 / 2}}\left(\frac{e x}{2 m+1}\right)^{(m+1 / 2)} \\
n_{m}(x) \rightarrow-\left(\frac{2}{(2 m+1) x}\right)^{1 / 2}\left(\frac{2 m+1}{e x}\right)^{(m+1 / 2)}
\end{array}\right\} \begin{aligned}
& \text { as } m \rightarrow \infty  \tag{7.1.36}\\
& \text { with } x \text { fixed }
\end{aligned}
$$

The qualitative description of these relations is that only the spherical Bessel function $j_{m}(x)$ is finite at zero argument, with all except the zero order approaching zero. Each of the other functions is singular at $x=0$, with an increasing growth rate as the order is increased. Also, replacing $x$ with $k r$ shows that the farfield behavior of $h_{m}^{(2)}(k r)$ matches that of a wave that propagates outward, whereas $h_{m}^{(1)}(k r)$ is associated with inward propagation. Furthermore, $j_{m}(k r)$ and $n_{m}(k r)$ represent standing wave patterns. The behavior at large orders is important because it affects the convergence of series that sums over all orders.

There are many formulas that relate the various spherical Bessel functions and their derivatives. A group that are particularly useful are the recurrence relations. The following uses $f_{m}(x)$ to denote $j_{m}(x), n_{m}(x), h_{m}^{(1)}(x)$, or $h_{m}^{(2)}(x)$,

$$
\begin{align*}
& f_{m-1}(x)+f_{m-1}(x)=\frac{2 m+1}{x} f_{m}(x), m>0 \\
& \frac{d}{d x} f_{m}(x)=\frac{m}{x} f_{m}(x)-f_{m+1}(x) \\
&=f_{m-1}(x)-\frac{m+1}{x} f_{m}(x), m>0 \\
&=\frac{1}{2 m+1}\left[m f_{m-1}(x)-(m+1) f_{m+1}(x)\right], m>0  \tag{7.1.37}\\
&\left(\frac{1}{x} \frac{d}{d x}\right)^{n}\left[x^{m+1} f_{m}(x)\right]=x^{m-n+1} f_{m-n}(x) \\
&\left(\frac{1}{x} \frac{d}{d x}\right)^{n}\left[x^{-m} f_{m}(x)\right]=(-1)^{n} x^{-m-n} f_{m+n}(x)
\end{align*}
$$

The first equation may be used to evaluate higher order functions in terms of lower order values. We will have frequent need for derivatives of Bessel functions, so it is good idea to create a subprogram that implements the first equation for $d f_{m}(x) / d x$. The last two relations seldom come into play in the analysis of a specific system. As is true for Legendre functions, the handbook compiled by Abramowitz and Stegun ${ }^{4}$ provides a compact resource for many other formulas. The book by Sneddon ${ }^{5}$ is a reference for those seeking detailed analyses of Bessel functions.

[^3]
### 7.1.2 Description of the Pressure Field

The separation of variables analysis began by seeking the conditions for which $P=$ $\mathcal{F}(r) \mathcal{G}(\psi)$ is a solution of the Helmholtz equation. This led to the recognition that $\mathcal{G}(\psi)$ must be a Legendre function of integer order $m$ and $\mathcal{F}(r)$ must be one of the types of spherical Bessel functions. Because the Helmholtz equation is homogeneous, the trial solution may be multiplied by a constant. Furthermore, the equation is linear, so a sum of solutions at various orders also is a solution.

Which type of spherical Bessel function we should use depends on the extent of the fluid's domain. We began with the stated interest of describing radiation into an infinite domain. The Sommerfeld radiation condition requires an outgoing wave, which means that we should use whichever function behaves as $e^{-i k r} / r$ when $r$ is large. Reference to Eq. (7.1.35) shows that this function is $h^{(2)}(k r)$. Thus, the axisymmetric pressure exterior to a sphere is described by

$$
\begin{equation*}
P=\sum_{m=0}^{\infty} B_{m} h_{m}^{(2)}(k r) P_{m}(\cos \psi), \quad r>a \tag{7.1.38}
\end{equation*}
$$

To specialize this expression to the farfield, we use the asymptotic representation of $h_{m}^{(2)}(k r)$ contained in Eq. (7.1.35), which leads to

$$
\begin{equation*}
P_{\mathrm{ff}}=\sum_{m=0}^{\infty} B_{m} P_{m}(\cos \psi) \frac{e^{-i k r}}{k r} e^{i(m+1) \pi / 2} \tag{7.1.39}
\end{equation*}
$$

The pressure in a spherical cavity is a topic in Chap. 10. As a preliminary to that study, we recall that the analysis of the radially symmetric field in Sect.6.3.3 began with a solution that contains inward and outwardly propagating waves. The corresponding form here would use both kinds of spherical Hankel function. The finiteness condition in the previous analysis led to the recognition that these waves must have complex amplitudes that sum to zero. A similar analysis would lead to the same conclusion here. According to Eq. (7.1.31), the sum of $h_{m}^{(2)}(k r)$ and $h_{m}^{(1)}(k r)$ gives a spherical Bessel function, while the difference gives a spherical Neumann function. The latter is singular at $r=0$, so the general solution for the axisymmetric pressure within a spherical cavity is

$$
\begin{equation*}
P=\sum_{m=0}^{\infty} B_{m} j_{m}(k r) P_{\mathrm{ff}}(\cos \psi), \quad 0 \leq r<a \tag{7.1.40}
\end{equation*}
$$

### 7.1.3 Arbitrary Spatial Dependence

The most general situation is a pressure field that is neither radially nor axially symmetric. For example, a single dipole and a longitudinal quadrupole generate fields that are axisymmetric, but a lateral quadrupole is not. An investigation of the most general spherical wave would pursue a separation of variables analysis of the Helmholtz equation by starting with

$$
\begin{equation*}
P=\mathcal{F}(r) \mathcal{G}(\psi) \Theta(\theta) \tag{7.1.41}
\end{equation*}
$$

Such an analysis would lead to the recognition that the azimuthal function $\Theta(\theta)$ must be harmonic. This fact is apparent on a physical basis, because circling the polar axis at a fixed radial distance and polar angle must return us to the point where we began. Thus, it must be that $P(r, \psi, \theta)=P(r, \psi, \theta+2 \pi)$. This is the condition for representing the azimuthal dependence in a Fourier series. At our option, we can take $\Theta$ to be a complex exponential, $e^{i n \theta}$, where $n$ is any positive or negative integer, or we may consider it to be a real harmonic by letting it be $\sin (n \theta)$ or $\cos (n \theta)$.

Regardless of which representation is selected, the fact that $\Theta(\theta)$ is a harmonic function whose argument is $n \theta$ reduces the task to a determination of the product $\mathcal{F}(r) \mathcal{G}(\psi)$. Application of the separation of variables technique to the Helmholtz equation leads to the recognition that the azimuthal harmonic number $n$ affects the equation for $\mathcal{G}$, but not the one for $\mathcal{F}$. The solution for $\mathcal{G}$ is found to be an associated Legendre function $P_{m}^{n}(\cos \psi)$, which is obtained from $n$ differentiations of an $m$ th degree Legendre polynomial, followed by the substitution of $\eta=\cos \psi$. The specific definition is

$$
\begin{equation*}
P_{m}^{n}(\cos \psi)=\left.\left[-\left(1-\eta^{2}\right)^{1 / 2}\right]^{m} \frac{d^{n}}{d \eta^{n}} P_{m}(\eta)\right|_{\eta=\cos \psi} \tag{7.1.42}
\end{equation*}
$$

Because $P_{m}(\eta)$ is an $m$ th degree polynomial, any derivative above the $m$ th degree will be identically zero, from which it follows that

$$
\begin{equation*}
P_{m}^{n}(\cos \psi)=0 \text { if } n>m \tag{7.1.43}
\end{equation*}
$$

The index $n$ is said to be the order of the associated Legendre function and $m$ is its degree.

Because the separation equation for $\mathcal{F}(r)$ does not depend on the azimuthal number $n$, it is the same alternative set of Bessel functions as in the axisymmetric case. Thus, for each polar number $m$, there are azimuthal harmonics from 0 to $m$. The pressure for each $m, n$ pair consists of a product of a spherical Bessel or Hankel function of order $n$ for the radial dependence, a $P_{m}^{n}(\cos \psi)$ function, and a harmonic function of $n \theta$. The range of $m$ is infinite, so the general solution for the pressure is a double sum. For example, in the case of radiation, with a complex Fourier series used to represent the azimuthal dependence, the solution would have the form

$$
\begin{equation*}
P=\frac{1}{2} \sum_{m=0}^{\infty} \sum_{n=-m}^{m} B_{m, n} h_{m}(k r) P_{m}^{n}(\cos \psi) e^{i n \theta} \tag{7.1.44}
\end{equation*}
$$

The factors $P_{m}^{n}(\cos \psi) e^{i n \theta}$ describe an arbitrary angular dependence; they constitute the full set of spherical harmonics.

Like the axisymmetric set, the general set of spherical harmonics are orthogonal when integrated over a sphere. Indeed, procedures that are used to analyze an axisymmetric field in terms of spherical harmonics are readily modified to handle arbitrary fields. Several books ${ }^{6}$ treat the mathematical aspects of a general spherical harmonic series expansion and its application to acoustical systems. If we were to study situations where there is no symmetry, we would not encounter much that is new from a phenomenological viewpoint. The primary reason for this is the aforementioned nature of the radial dependence, which is independent of the azimuthal harmonic number.

EXAMPLE 7.2 Consider an axisymmetric field consisting of a monopole, a dipole, and a longitudinal quadrupole. This combination of sources is collocated, and the axes of the dipole and quadrupole are aligned. What is the spherical harmonic representation of the farfield?

## Significance

Experience in working with spherical harmonics is the primary emphasis. The result will highlight the different pictures of the farfield that results from a spherical harmonic series, rather than a multipole expansion.

## Solution

The multipole representation of the farfield pressure is provided by Eq. (6.5.142). To use it, we will let $z \equiv x_{3}$ be the axis of symmetry, and set the dipole such that its moment vector is aligned along $z$. A longitudinal quadrupole that is aligned in this direction has $Q_{3,3}$ as the only nonzero component. The spherical coordinate transformation gives $z=r \cos \psi$, so the given pressure is

$$
\begin{equation*}
P=\left[A+i k D_{3} \cos \psi-k^{2} Q_{3,3}(\cos \psi)^{2}\right] \frac{e^{-i k r}}{r} \tag{1}
\end{equation*}
$$

where the strength factors are unspecified.
The spherical harmonic description of an axisymmetric field is Eq.(7.1.38), and Eq. (7.1.39) is the corresponding farfield approximation. The $B_{m}$ coefficients are found by applying the orthogonality property. It is somewhat easier to formulate the

[^4]analysis in terms of $\eta=\cos \psi$. The actual and series descriptions of $P$ are equated, and both sides are multiplied by $P_{j}(\eta)$. Integration over $-1 \leq \eta \leq 1$ leads to
\[

$$
\begin{aligned}
\sum_{m=0}^{\infty} \int_{-1}^{1}\left[B_{m} P_{m}(\eta)\right. & \left.\frac{e^{-i k r}}{k r} e^{i(m+1) \pi / 2}\right] P_{j}(\eta) d \eta \\
& =\int_{-1}^{1}\left[A+i k D_{3} \eta-k^{2} Q_{3,3} \eta^{2}\right] \frac{e^{-i k r}}{r} P_{j}(\eta) d \eta
\end{aligned}
$$
\]

All terms in the summation integrate to zero, except for the term for which $m=j$. The result is

$$
\begin{equation*}
B_{j}=\left(\frac{2 j+1}{2}\right) k e^{-i(j+1) \pi / 2} \int_{-1}^{1}\left[A+i k D_{3} \eta-k^{2} Q_{3,3} \eta^{2}\right] P_{j}(\eta) d \eta \tag{2}
\end{equation*}
$$

We could evaluate the integral numerically as we did in the previous example. However, the facts that $P_{j}(\eta)$ is a polynomial in $\eta$, and that it is multiplied by powers of $\eta$, suggest we try analytical integration. One approach would be to work with the series representation of Legendre functions, Eq. (7.1.11), but a simpler analysis is available. Equation (7.1.12) indicates that $P_{0}(\eta)=1, P_{1}(\eta)=\eta$, and [2 $\left.P_{2}(\eta)+P_{0}(\eta)\right] / 3=\eta^{2}$, so Eq. (2) may be written as

$$
\begin{aligned}
B_{j}= & \left(\frac{2 j+1}{2}\right) k e^{-i(j+1) \pi / 2} \int_{-1}^{1}\left[A P_{0}(\eta)+i k D_{3} P_{1}(\eta)\right. \\
& \left.-k^{2} Q_{3,3} \frac{2 P_{2}(\eta)+P_{0}(\eta)}{3}\right] P_{j}(\eta) d \eta
\end{aligned}
$$

The orthogonality property describes the integral of each term. Specifically, $P_{n}(\eta)$ $P_{j}(\eta)$ integrates to zero unless $j=n$, in which case the integral is $2 /(2 n+1)$. Thus, the result is that

$$
B_{0}=-i k A+i k^{3} Q_{33}, \quad B_{1}=-i k^{2} D_{3}, \quad B_{2}=-(2 / 3) i Q_{3,3}
$$

The most notable attribute is that the spherical harmonics are not in one-to-one correspondence to the point source distributions. As the source distribution goes up in order, more harmonics are required to represent it. In general, the field of an axisymmetric $n$-pole will be proportional to $\cos \psi$ raised to powers up to $n$. The resulting $B_{j}$ coefficients for $j \leq n$ will contain a contribution from that pole if $j$ and $n$ have the same parity.

Given that several spherical harmonics are required to describe the higher multipoles, the utility of spherical harmonics might seem to be questionable. This is a limited view. The multipole construction requires that the sources be a compact set, whereas spherical harmonics may be used to describe any spherical field. Furthermore, the orthogonality property enables a complicated field to be mapped into a simpler representation as a series.

### 7.2 Radiation from a Spherical Body

### 7.2.1 Analysis

Determination of the field radiated by an axisymmetrically vibrating sphere is our first application of a spherical harmonic series. The axis of symmetry is designated as the polar axis $z$. The radial vibration of the sphere is an arbitrary function of the polar angle, $V(\psi)$. The sphere, whose radius is $a$, is surrounded by an ideal fluid. Cases in which the sphere executes a breathing mode vibration, or oscillates as a rigid body, which we studied in Chap. 6, are special cases. Generalization to an arbitrary vibration pattern will lead to insights about what aspects of the vibration pattern are important to the pressure field and radiated power.

In addition to the spherical harmonic series for the pressure in a radiation problem, Eq.(7.1.38), we will require the associated radial velocity. It is obtained by applying Euler's equation to the pressure field. These expressions are

$$
\begin{align*}
P & =\sum_{m=0}^{\infty} B_{m} h_{m}(k r) P_{m}(\cos \psi)  \tag{7.2.1}\\
V_{r} & =\frac{i}{\rho_{0} c} \sum_{m=0}^{\infty} B_{m} h_{m}^{\prime}(k r) P_{m}(\cos \psi)
\end{align*}
$$

Explicit designation of the Hankel function as being the second kind has been abandoned because that is the only kind that is relevant to the analysis. (If we had adopted the $\exp (-i \omega t)$ convention for a harmonic function, it would use the first kind of Hankel function.) Also, the prime designates differentiation of the function with respect to its argument, that is,

$$
\begin{equation*}
\left.h_{m}^{\prime}(k r) \equiv \frac{d}{d \xi} h_{m}^{(2)}(\xi)\right|_{\xi=k r} \tag{7.2.2}
\end{equation*}
$$

The preceding is descriptive of any axisymmetric field. To fit it to the present radiation problem, we match the radial particle velocity at the sphere's surface to that of the sphere. Toward that end, the surface velocity is represented as a spherical harmonic series,

$$
\begin{equation*}
V(\psi)=v_{0} \sum_{m=0}^{\infty} V_{m} P_{m}(\cos \psi) \tag{7.2.3}
\end{equation*}
$$

The parameter $v_{0}$ is a measure of the magnitude of the surface velocity, such as its maximum. It has been factored out of the series in order that the $V_{m}$ coefficients be dimensionless. An expression for these coefficients is found by invoking the orthogonality property of the spherical harmonics, which gives

$$
\begin{align*}
V_{m} & =\left(m+\frac{1}{2}\right) \int_{0}^{\pi} \frac{V(\psi)}{v_{0}} P_{m}(\cos \psi) \sin \psi d \psi  \tag{7.2.4}\\
& =\left(m+\frac{1}{2}\right) \int_{-1}^{1} \frac{V\left(\cos ^{-1}(\eta)\right)}{v_{0}} P_{m}(\eta) d \eta
\end{align*}
$$

Which of these forms is more conducive to the evaluation of the integral depends on the nature of the $V(\psi)$ function.

The continuity condition requires that $V_{r}=V(\psi)$ at $r=a$. We perform this matching by using Eq. (7.2.3) to represent the sphere's motion and Eq. (7.2.1) to represent the response of the fluid. Thus, we have

$$
\begin{equation*}
\frac{i}{\rho_{0} c} \sum_{m=0}^{\infty} B_{m} h_{m}^{\prime}(k a) P_{m}(\cos \psi)=v_{0} \sum_{m=0}^{\infty} V_{m} P_{m}(\cos \psi) \tag{7.2.5}
\end{equation*}
$$

The Legendre functions constitute an orthogonal set, so the preceding relation must be satisfied separately by each term. This leads to

$$
\begin{equation*}
B_{m}=\rho_{0} c v_{0} \frac{V_{m}}{i h_{m}^{\prime}(k a)} \tag{7.2.6}
\end{equation*}
$$

The resulting expressions for particle velocity and pressure are

$$
\begin{gather*}
P=\rho_{0} c v_{0} \sum_{m=0}^{\infty} V_{m} \frac{h_{m}(k r)}{i h_{m}^{\prime}(k a)} P_{m}(\cos \psi) \\
V_{r}=v_{0} \sum_{m=0}^{\infty} V_{m} \frac{h_{m}^{\prime}(k r)}{h_{m}^{\prime}(k a)} P_{m}(\cos \psi) \tag{7.2.7}
\end{gather*}
$$

Although these expressions call for infinite summations, in most cases, adequate accuracy will be obtained when the summations are truncated at a relatively small number of terms.

It is useful to verify that these expressions are consistent with prior developments. In the case of a sphere that executes a breathing mode vibration, the surface velocity $V(\psi)=v_{0}$. Because $P_{0}(\cos \psi)=1$, the only nonzero velocity coefficient is $V_{0}=1$. From Eq. (7.1.28), we have

$$
\begin{equation*}
h_{0}(x)=\frac{i e^{-i x}}{x} \Longrightarrow h_{0}^{\prime}(x)=\frac{x-i}{x^{2}} e^{-i x} \tag{7.2.8}
\end{equation*}
$$

These expressions, evaluated at $x=k r$ for the former and $x=k a$ for the latter, are substituted into the $m=0$ term in the summation, which yields

$$
\begin{equation*}
P=\rho_{0} c v_{0} \frac{i e^{-i k r} / r}{i\left[\frac{k a-i}{(k a)^{2}}\right] e^{-i k a}} \tag{7.2.9}
\end{equation*}
$$

When this expression is simplified, the result is the same as the first of Eq. (6.3.7).
Proof that the spherical harmonic description reduces to the previous result for a translating sphere is a little more complicated. In this case, $V(\psi)=v_{0} \cos \psi$, so $V_{1}=1$ is the sole nonzero coefficient. According to Eq.(7.1.28), the first-order Hankel function and its derivative are

$$
\begin{equation*}
h_{1}(x)=\left(\frac{-x+i}{x^{2}}\right) e^{-i x} \Longrightarrow h_{1}^{\prime}(x)=\left(\frac{i x^{2}+2 x-2 i}{x^{3}}\right) e^{-i x} \tag{7.2.10}
\end{equation*}
$$

Substitution of these terms into the $m=1$ term in the summation gives

$$
\begin{equation*}
P=\rho_{0} c v_{0} \frac{\left[\frac{-k r+i}{(k r)^{2}}\right] e^{-i k r}}{i\left[\frac{i(k a)^{2}+2 k a-2 i}{(k a)^{3}}\right]} \tag{7.2.11}
\end{equation*}
$$

Simplification of this expression leads to Eq. (6.5.102).

### 7.2.2 Important Limits

Not much about the nature of the field is apparent in the general solution, so we turn to asymptotic trends, beginning with the farfield. When $k r$ is large, Eq. (7.1.35) indicates that

$$
\begin{equation*}
h_{m}(k r) \rightarrow \frac{e^{-i k r}}{k r} e^{i(m+1) \pi / 2}, \quad h_{m}^{\prime}(k r) \rightarrow-i h_{m}(k r) \text { as } k r \rightarrow \infty \tag{7.2.12}
\end{equation*}
$$

Application of these trends to Eq. (7.2.7) leads to

$$
\begin{equation*}
P_{\mathrm{ff}}=\rho_{0} c\left(V_{r}\right)_{\mathrm{ff}}=\rho_{0} c v_{0} \frac{e^{-i k r}}{k r} \sum_{m=0}^{\infty} V_{m} \frac{e^{i m \pi / 2}}{h_{m}^{\prime}(k a)} P_{m}(\cos \psi) \tag{7.2.13}
\end{equation*}
$$

Typically, the farfield property of interest is the directivity. Previously, we defined $\mathcal{D}(\psi)$ such that its maximum value is one. It is more convenient here to use as the reference pressure $\rho_{0} c v_{0}(a / r)$. Only the pressure magnitude is considered. Thus, we have

$$
\begin{equation*}
\mathcal{D}(\psi) \equiv\left|\frac{P_{\mathrm{ff}}}{\rho_{0} c v_{0}}\right|\left(\frac{r}{a}\right)=\frac{1}{k a}\left|\sum_{m} \frac{V_{m} e^{i m \pi / 2}}{h_{m}^{\prime}(k a)} P_{m}(\cos \psi)\right| \tag{7.2.14}
\end{equation*}
$$

Another use of the farfield pressure is to evaluate the radiated power. Toward that end, we form the time-averaged intensity at large $k r$, which is $\left(I_{r}\right)_{\mathrm{av}}=P_{\mathrm{ff}} P_{\mathrm{ff}}^{*} /\left(2 \rho_{0} c\right)$. To evaluate the product, we change the summation index for one factor, which leads to

$$
\begin{equation*}
\left(I_{r}\right)_{\mathrm{av}}=\frac{1}{2} \rho_{0} c v_{0}^{2}\left(\frac{1}{k r}\right)^{2} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} V_{j} V_{m}^{*} \frac{e^{i(j-m) \pi / 2}}{h_{j}^{\prime}(k a)\left(h_{m}^{\prime}(k a)\right)^{*}} P_{j}(\cos \psi) P_{m}(\cos \psi) \tag{7.2.15}
\end{equation*}
$$

The radiated power is found by integrating this expression over the surface of a very large sphere whose radius is $r$. Because of the axisymmetry of the field, a suitable differential surface element is $d \mathcal{S}=2 \pi r \sin \psi(r d \psi)$. When we bring the surface integral inside the double sum, we find that

$$
\begin{equation*}
\mathcal{P}=\frac{\pi \rho_{0} c}{k^{2}} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} V_{j} V_{m}^{*} \frac{e^{i(j-m) \pi / 2}}{h_{j}^{\prime}(k a)\left(h_{m}^{\prime}(k a)\right)^{*}} \int_{0}^{\pi} P_{j}(\cos \psi) P_{m}(\cos \psi) \sin \psi d \psi \tag{7.2.16}
\end{equation*}
$$

The integral is merely the orthogonality condition for Legendre functions, which means that all terms in the double sum for which $j \neq m$ vanish. This simplifies the result to

$$
\begin{equation*}
\mathcal{P}=\pi \rho_{0} c v_{0}^{2} a^{2} \frac{1}{(k a)^{2}} \sum_{m=0}^{\infty} \frac{2}{2 m+1} \frac{\left|V_{m}\right|^{2}}{\left|h_{m}^{\prime}(k a)\right|^{2}} \tag{7.2.17}
\end{equation*}
$$

This expression exhibits an important property: The spherical harmonics of the surface velocity are uncoupled in the radiated power. In other words, each harmonic gives a fixed contribution to the radiated power, independently of the magnitude of the other harmonics.

Another important feature is related to convergence properties. When the harmonic number $m$ becomes increasingly large, $h_{m}^{\prime}(k a)$ will increase rapidly, see Eq. (7.1.36). Consequently, the high-order terms in the series for the radiated power have much less importance than they do for the surface velocity. Recall that increasing spherical harmonic number represents more "wiggles" in the spatial fluctuation. It follows that this behavior is another manifestation of the general notion that waves on a surface that have a short wavelength tend to not radiate to the farfield.

Another limit of interest is a small sphere. The expressions for pressure and radial velocity depend on the radius nondimensionally as $k a$, so the small radius limit also is the low-frequency limit. The asymptotic behavior of $h_{m}^{\prime}(k a)$ may be found by differentiating $h_{m}^{(2)}(x)$ in Eq. (7.1.34), which gives

$$
\begin{equation*}
h_{m}^{\prime}(k a) \rightarrow-4 i \frac{(m+1)(2 m)!}{(2 k a)^{m+2}} \text { as } k a \rightarrow 0 \tag{7.2.18}
\end{equation*}
$$

With these, the pressure in Eq. (7.2.7) becomes

$$
\begin{equation*}
P=\rho_{0} c v_{0} \sum_{m=0}^{\infty} V_{m} \frac{(2 k a)^{m+2}}{4(m+1)(2 m)!} h_{m}(k r) P_{m}(\cos \psi), \quad k a \ll 1 \tag{7.2.19}
\end{equation*}
$$

Before we interpret this expression, let us consider the special case of a radially symmetric vibration. The only nonzero velocity coefficient is $V_{0}=V(\psi) / v_{0}=1$. The definitions are $h_{0}(k r) \equiv i e^{-i k r} /(k r)$ and $P_{0}(\cos \psi)=1$, and we also have

$$
\begin{equation*}
\left.h_{0}^{\prime}(k a) \equiv \frac{d}{d \xi}\left(\frac{i e^{-i \xi}}{\xi}\right)\right|_{\xi=k a} \equiv\left(\frac{k a-i}{k^{2} a^{2}}\right) e^{-i k a} \tag{7.2.20}
\end{equation*}
$$

When $V_{0}$ is the only nonzero coefficient, Eq. (7.2.9) for arbitrary values of $k r$ and $k a$ gives

$$
\begin{equation*}
P=\rho_{0} c v_{0} \frac{h_{0}(k r)}{i h_{0}^{\prime}(k a)}=\rho_{0} c v_{0}\left(\frac{k a}{k a-i}\right) \frac{a}{r} e^{-i k(r-a)} \tag{7.2.21}
\end{equation*}
$$

Now consider Eq.(7.2.19) when $k a \ll 1$. Each term in the summation raises $k a$ to a higher power, so it follows that the dominant term corresponds to the lowest harmonic $m$ for which $V_{m}$ is nonzero. Usually, this corresponds to $m=0$, in which case, this description reduces to

$$
\begin{equation*}
P=\rho_{0} c v_{0}(k a)^{2} h_{0}(k r) P_{0}(\cos \psi) \equiv \rho_{0} c v_{0}(i k a) \frac{a}{r} e^{-i k r} \tag{7.2.22}
\end{equation*}
$$

Because $e^{i k a} \approx 1$ when $k a \ll 1$, Eq. (7.2.21) reduces to this expression. Hence, we conclude that in the small sphere/low-frequency limit, the only aspect of the surface motion that is important is the volume velocity, $\hat{Q}=4 \pi a^{2} v_{0}$. Only if the volume velocity is zero does the pattern of the surface vibration affect the acoustic field.

The high-frequency/large sphere limit also leads to an interesting observation. Let $M$ denote the lowest spherical harmonic at which the summation in Eq. (7.2.3) may be truncated without reducing the accuracy below the desired value. The definition of high frequency is that $k a \gg M$. In that case, an expression for $h_{m}^{\prime}(k a)$ at all significant orders $m$ may be obtained by differentiation of the large argument approximation of $h_{m}^{(2)}(k a)$, Eq. (7.1.35), which gives

$$
\begin{equation*}
h_{m}^{\prime}(k a) \rightarrow \frac{-i}{k a} e^{-i(k a-(m+1) \pi / 2)} \text { as } k a \rightarrow \infty \tag{7.2.23}
\end{equation*}
$$

This approximation converts Eq. (7.2.13) to

$$
\begin{equation*}
P_{\mathrm{ff}} \rightarrow \rho_{0} c v_{0} \frac{a}{r} e^{-i k(r-a)} \sum_{m=0}^{M} V_{m} P_{m}(\cos \psi) \tag{7.2.24}
\end{equation*}
$$

The sum is the truncated representation of $V(\psi) / v_{0}$, so the corresponding directivity is $\mathcal{D}=|V(\psi)| / v_{0}$. In other words, the directivity replicates the pattern of the sphere's radial velocity pattern.

This limiting behavior is readily explained. Because $P_{m}(\eta)$ is an $m$ th degree polynomial, it follows that $P_{m}(\cos \psi)$ may be represented as a Fourier series in $\psi$ whose highest harmonic is $\cos (m \psi)$. The wavelength of a surface vibration on the sphere at this harmonic is the circumference divided by $m, \lambda_{m}=2 \pi a / m$. The shortest wavelength that is significant to the vibration is $\lambda_{M}$. The acoustic wavelength is $2 \pi / k$. Thus, defining high frequency to be $k a \gg M$ is equivalent to stating that the high-frequency regime is such that all significant $\lambda_{m}$ are much greater than $2 \pi / k$. In other words, spherical harmonics whose surface wavelength is much larger than the acoustic wavelength are responsible for most of the radiation to the farfield. This is reminiscent of the behavior that was observed for surface waves on an infinite planar surface, see Sect. 5.1.

EXAMPLE 7.3 A spherical cap is the region subtending an angle $\psi_{0}$ relative to a radial line that is designated as the $z$-axis. This cap executes a translational vibration $\operatorname{Re}\left(v_{0} e^{i \omega t}\right)$ parallel to $z$-axis, whereas the remainder of the sphere is rigid and stationary. The corresponding complex amplitude of the radial velocity on the surface $r=a$ is $V_{r}=v_{\text {cap }} \bar{e}_{z} \cdot \bar{e}_{r}=v_{0} \cos \psi$ if $0 \leq \psi \leq \psi_{0}, V_{r}=0$ if $\psi_{0}<\psi \leq \pi$. Derive expressions for the coefficients $V_{m}$ of the spherical harmonic series. Then, evaluate these coefficients for the case where $\psi_{0}=30^{\circ}$, and identify the harmonic number $N$ for which a reconstruction of $V_{r}$ deviates by no more than $0.2 v_{0}$ from the actual $V_{r}$. Use this value of $N$ to evaluate the pressure along the $z$-axis, $\psi=0$, for $k a=12$. Use that result to estimate the value of $k r$ at which the farfield begins. Then, use the same value of $N$ to evaluate the farfield directivity, and compare that result to the high-frequency asymptotic trend.

## Significance

A simplified version of this problem has appeared in many books. The simplification entails considering the velocity of the spherical cap to be radial, rather than being proportional to $\cos \psi$. The availability of mathematical software, coupled with a little ingenuity regarding the application of the recurrence relations for Legendre polynomials, makes it possible to carry out a consistent analysis. The outcome will show the utility of some of the asymptotic properties of the pressure field.

## Solution

We begin by evaluating the $V_{m}$ coefficients, which are given by Eq. (7.2.4). The function of interest here is zero beyond $\psi=\psi_{0}$, so we have

$$
\begin{align*}
V_{m} & =\left(m+\frac{1}{2}\right) \int_{0}^{\psi_{0}} \cos \psi P_{m}(\cos \psi) \sin \psi d \psi \\
& =\left(m+\frac{1}{2}\right) \int_{\cos \psi_{0}}^{1} \eta P_{m}(\eta) d \eta \tag{1}
\end{align*}
$$

The question now is how to evaluate the integral? We could use a numerical integration algorithm. If we were to follow that route, we would encounter difficulty because a rather large number of terms are required to meet the imposed convergence criterion, and $P_{m}(\eta)$ oscillates rapidly when $m$ is large. An alternative is to replace $P_{m}(\eta)$ with its series representation. Here too, large $m$ would lead to difficulty because the factorials in the numerator and denominator of the series for $P_{m}(\eta)$ would be very large. Our approach relies on the recurrence relations. We use the first of Eq. (7.1.14) to replace $\eta P_{m}(\eta)$ with term that are purely Legendre polynomials. To that result, we will apply the third of Eq.(7.1.14) in order to replace those polynomials with derivatives of Legendre polynomials. These derivatives may be integrated directly. The operations are

$$
\begin{aligned}
V_{m}= & \left(m+\frac{1}{2}\right) \int_{\cos \psi_{0}}^{1} \frac{(m+1) P_{m+1}(\eta)+m P_{m-1}(\eta)}{2 m+1} d \eta \\
= & \frac{1}{2} \int_{\cos \psi_{0}}^{1}\left\{\frac{m+1}{2 m+3}\left[\frac{d}{d \eta} P_{m+2}(\eta)-\frac{d}{d \eta} P_{m}(\eta)\right]\right. \\
& \left.+\frac{m}{2 m-1}\left[\frac{d}{d \eta} P_{m}(\eta)-\frac{d}{d \eta} P_{m-2}(\eta)\right]\right\} d \eta
\end{aligned}
$$

All Legendre polynomials are unity at $\eta=1$, so we find that

$$
\begin{align*}
V_{m}= & -\frac{1}{2}\left[\frac{m+1}{2 m+3} P_{m+2}\left(\cos \psi_{0}\right)+\frac{(2 m+1)}{(2 m+3)(2 m-1)} P_{m}\left(\cos \psi_{0}\right)\right. \\
& \left.-\frac{m}{2 m-1} P_{m-2}\left(\cos \psi_{0}\right)\right] \tag{2}
\end{align*}
$$

These relations are valid for $m \geq 2$. The first two coefficients may be found directly by substituting $P_{0}(\eta) \equiv 1$ and $P_{1}(\eta) \equiv \eta$. The result is

$$
\begin{equation*}
V_{0}=\frac{1}{4}\left[1-\left(\cos \psi_{0}\right)^{2}\right], \quad V_{1}=\frac{1}{2}\left[1-\left(\cos \psi_{0}\right)^{3}\right] \tag{3}
\end{equation*}
$$

For comparison, the simplified version, in which $V_{r}=v_{0}$, would not have the $\eta$ factor in the integrand. The third of the recurrence relations in Eq.(7.1.14) would suffice in that case, which would lead to

$$
\begin{align*}
&\left(V_{m}\right)_{\operatorname{simp}}=\left(m+\frac{1}{2}\right) \int_{\cos \psi_{0}}^{1} P_{m}(\eta) d \eta \\
&=-\frac{1}{2}\left[P_{m+1}\left(\cos \psi_{0}\right)-P_{m-1}\left(\cos \psi_{0}\right)\right], \quad m \geq 1  \tag{4}\\
& \quad\left(V_{0}\right)_{\operatorname{simp}}=\frac{1}{2}\left[1-\left(\cos \psi_{0}\right)\right]
\end{align*}
$$

The consistent and simplified $V_{m}$ coefficients are described by Fig. 1. The two sets of values are quite close. This is a consequence of the fact that for $0 \leq \psi \leq 30^{\circ}$, the value of $\cos \psi$ is close to unity. The discrepancy would grow if $\psi_{0}$ were larger.


Figure 1.

The $V_{m}$ values decrease with increasing $m$, but slowly and in an oscillatory manner. As we add more terms to the series for $V(\psi)$, sometimes the additional terms will add, and sometimes they will subtract. Consequently, examination of the magnitude of the coefficients gives an uncertain picture of the significance of each term in the series. We must assess convergence by constructing $V(\psi)$ according to Eq. (7.2.3). Let $V(\psi, N)$ denote the value obtained when the series is truncated at $m=N$. The error metric of interest is $\left|V(\psi, N)-\cos (\psi)\left[h(\psi)-h\left(\psi-\psi_{0}\right)\right]\right|$. The following table lists the maximum error for $0 \leq \psi \leq \pi$ for several $N$.

$$
\begin{array}{c|c|c|c|c|c}
N & 20 & 40 & 60 & 80 & 100 \\
\hline \text { Error (\%) } & 37 & 32 & 25 & 21 & 16
\end{array}
$$

The error at $N=80$ is very close to the requested value $0.2 v_{0}$, so that is what we will use for the remainder of this investigation.

Figure 2 depicts the reconstruction of $V(\psi) / v_{0}$ obtained for two values of $N$. The synthesis of Eq. (7.2.3) can be done with the matrix algorithm described in Example 7.1. Let $\{V\}$ be the column vector of the $N+1$ velocity coefficients $V_{m}$, and let $\{\psi\}$ be a set of $J$ polar angles at which it is desired to evaluate the series, so that $\{\eta\}=\cos (\{\psi\})$. A rectangular $J \times(N+1)$ array $[L]$ is defined such that column $m$ is $P_{m}(\{\eta\})$ for $m=0,1, \ldots, N$. Then, the radial velocity on the surface is given by $\left\{V_{\text {surf }}\right\}=[L]\{V\}$. The result for $N=20$ exhibits a good resemblance of the actual function, while $N=80$ is very close, with a minor Gibb's phenomenon at $\psi=\psi_{0}=30^{\circ}$.


Figure 2.

The implication thus far is that a good representation of the surface velocity is necessary to describe the pressure field adequately, but this is not so. Let us rewrite the solution for pressure in Eq. (7.2.7) as

$$
P_{\mathrm{ff}}=\rho_{0} c v_{0} \sum_{m=0}^{\infty} \Gamma_{m} h_{m}(k r) P_{m}(\cos \psi), \quad \Gamma_{m}=\frac{V_{m}}{i h_{m}^{\prime}(k a)}
$$

Correspondingly, Eq. (7.2.14) for the directivity and Eq. (7.2.17) for the radiated power become

$$
\begin{aligned}
\mathcal{D}(\psi) & =\frac{1}{k a}\left|\sum_{m=0}^{\infty} \Gamma_{m} e^{i m \pi / 2} P_{m}(\cos \psi)\right| \\
\mathcal{P} & =\pi \rho_{0} a^{2} c v_{0}^{2} \sum_{m=0}^{\infty} \frac{2}{2 m+1} \frac{\left|\Gamma_{m}\right|^{2}}{(k a)^{2}}
\end{aligned}
$$

The $\Gamma_{m}$ values indicate which spherical harmonics are important to the pressure field. Figure 3 displays the $\Gamma_{m}$ values and the magnitude of each term in the series for radiated power. These graphs would be unaltered if we truncated the series for $V(\psi)$ at $N=20$, so it is evident that the small scale details of the surface motion are unimportant for the radiated field.


Figure 3.

The next feature to examine is the farfield and where it begins. The primary aspect of the farfield in any case is that the pressure decays as $1 / r$ along any radial line. Thus, a plot of $k r\left|P /\left(\rho_{0} c v_{0}\right)\right|$ as a function of $k r$ will approach a constant value as the $k r$ reaches into the farfield. The only radial line we were requested to consider is $\psi=0$, but one should carry out such computations along several lines in order to assure that what is observed on a specific line is not anomalous. An evaluation of
$|P|$ as a function of $k r$ also may be carried out by the matrix algorithm. Specifically, $\{k r\}$ is a column vector of $k r_{j}$ values, and column $m$ of [ $h$ ] holds the values of $h_{m}(\{k r\})$ for $m=0, \ldots, N$, and $[L]$ is a diagonal $N+1$ array whose elements are the $P_{m}(\cos \{\psi\})$ values. Then, the values of pressure at locations $k r_{j}$ along the radial line at fixed $\psi$ are $[h][L]\{V\}$.

The center line is $\psi=0$, so all of the Legendre function values are one. The result of this computation is graphed in Fig.4. The onset of a constant value of $k r P$ is gradual, but it seems that a good estimate would be that the product approaches a constant value in the vicinity of $k r=100$. The value of $k a$ is 12 , so the observation that the farfield begins at $k r=100$ suggests that the farfield begins at a $k r$ value whose order of magnitude is $(k a)^{2}$. This turns out to be a good empirical rule in general. Specifically, if $a$ is a measure of the size of an object that radiates an acoustic signal, then the farfield will begin at a nondimensional distance $k r_{\mathrm{ff}}=\beta(k a)^{2}$, where $\beta$ is a value having unit order of magnitude. The identified value of $\beta$ for the present system is approximately 0.75 .


Figure 4.

The last property to consider is the farfield directivity. From Eq. (7.2.14), we have

$$
\mathcal{D}(\psi)=\frac{1}{k a}\left|\sum_{m=0}^{\infty} \Gamma_{m} e^{i(m+1) \pi / 2} P_{m}(\cos \psi)\right|
$$

The matrix-based algorithm for evaluating a series may also be implemented here. The directivity obtained from the consistent and simplified analyses appear in Fig. 5. To avoid obscuring details, the data has been displayed as a rectangular plot, rather than the standard polar plot. It is not surprising, given the closeness of the two sets of coefficients, that the directivity obtained from the simplified representation of $V(\psi)$ is quite good. We can see a hint of the high-frequency limit, in which the directivity is proportional to the surface velocity. However, the value at $\psi=0$ is $40 \%$ greater than the limiting value, and $\mathcal{D}$ does not drop abruptly from $\cos \psi_{0}$ to zero at $\psi_{0}$.


Figure 5.

### 7.2.3 Symmetry Plane

Anyone who is interested in sound reproduction systems is aware that speakers usually are mounted on a baffle. A baffle is a boundary, usually flat, whose purpose is to direct the radiated sound to one side. A system that features a sphere in a baffle is a hemispherical dome tweeter, which is commonly used in speaker systems to create and radiate the higher frequencies in the audible spectrum. In this arrangement, a circular cap is mounted on the baffle, but is free to vibrate. In practice, the cap might not be a full hemisphere, but we will employ that idealization because it has an analytical solution. The example will compare the radiated field obtained with and without the baffle.

A hemispherical dome tweeter is not the only configuration in which a baffle guides spherical waves. A great simplification of a surface ship represents it as a sphere. Suppose that the sphere's buoyancy is such that its equator coincides with the free surface. The characteristic impedance of water is much greater than that of air. To a wave in the water, the surface appears to be pressure-release, whereas it appears to be rigid to an acoustic wave in the air. Thus, the model for acoustic radiation from the (idealized) ship into the air would consider a hemisphere in an infinite, planar, rigid baffle. In contrast, the model for acoustic radiation from the ship into the water would consider a hemisphere in an infinite, planar, pressure-release baffle.

Figure 7.3 describes the geometry of a hemisphere in an infinite baffle. The portion of the boundary that is the hemisphere is defined by $r=a, 0 \leq \psi<\pi / 2$, and the baffle is defined by $r>a, \psi=\pi / 2$. Velocity continuity requires that the particle velocity matches the hemisphere's radial velocity. In addition, if the baffle is rigid, the particle velocity normal to the baffle must vanish. Because $\bar{e}_{r}$ is parallel to the baffle and $\bar{e}_{\psi}$ is perpendicular to it, this requirement is that $\bar{V} \cdot \bar{e}_{\psi}=0$ at $\psi=\pi / 2$ for $r>a$. The case of a pressure-release baffle requires that $P=0$ at the same locations. Thus, the problem to solve is

Fig. 7.3 A hemisphere mounted on an infinite planar baffle


$$
\left\{\begin{array}{c}
\nabla^{2} P+k^{2} P=0, \quad r>a, 0 \leq \psi<\pi / 2  \tag{7.2.25}\\
\frac{\partial P}{\partial r}=-i \omega \rho_{0} V_{h}, \quad r=a, 0 \leq \psi<\pi / 2 \\
\text { Rigid baffle: } \frac{\partial P}{\partial \psi}=0 \\
\text { Pressure-release baffle: } P=0
\end{array}\right\}, r>a, \psi=\pi / 2
$$

where $V_{h}$ is the complex amplitude of the radial velocity of the hemisphere.
The response with either type of baffle may be obtained as a special case of a general axisymmetric wave. To see how to do so, we recall the method of images. If the boundary of an infinite half-space is rigid, the field may be obtained by placing a mirror image(s) behind the boundary of the actual source(s), with all images vibrating in-phase with their matching actual source. Here, the image of a hemisphere is also hemisphere, so the source and its image form a full sphere. If $V_{h}(\psi)$ is the radial velocity at location $\psi$ on the hemisphere, then the mirror image point is at $\pi-\psi$. Hence, the radial velocity distribution for the image sphere is

$$
\text { Rigid baffle: } V(\psi)=\left\{\begin{array}{l}
V_{h}(\psi), 0 \leq \psi<\pi / 2  \tag{7.2.26}\\
V_{h}(\pi-\psi), \pi / 2 \leq \psi<\pi
\end{array}\right.
$$

In the case where the boundary of a half-space is pressure-release, the mirror image hemisphere vibrates $180^{\circ}$ out-of-phase relative to the source hemisphere. Thus, the radial velocity of the full sphere in this case is

$$
\text { Pressure-release baffle: } V(\psi)=\left\{\begin{array}{l}
V_{h}(\psi), 0 \leq \psi<\pi / 2  \tag{7.2.27}\\
-V_{h}(\pi-\psi), \pi / 2 \leq \psi<\pi
\end{array}\right.
$$

The acoustic field that is radiated must share the same symmetry properties as the excitation. Thus, if the baffle is rigid, the pressure must be an even function of $\psi$ relative to $\psi=\pi / 2$, in other words, $P(r, \pi-\psi)=P(r, \psi)$. This requires that the only nonzero spherical harmonics in the general solution, Eq. (7.2.1), be those whose degree is even, $P_{0}(\cos \psi), P_{2}(\cos \psi), \ldots$. Thus, the method of images tells us that the field radiated by a hemisphere mounted on an infinite rigid baffle must be describable as

$$
\begin{gather*}
P=\rho_{0} c v_{0} \sum_{m=0}^{\infty} V_{2 m} \frac{h_{2 m}(k r)}{i h_{2 m}^{\prime}(k a)} P_{2 m}(\cos \psi)  \tag{7.2.28}\\
V_{r}=v_{0} \sum_{m=0}^{\infty} V_{2 m} \frac{h_{2 m}^{\prime}(k r)}{h_{2 m}^{\prime}(k a)} P_{2 m}(\cos \psi)
\end{gather*}
$$

To verify the correctness of this ansatz, we evaluate $V_{\psi}$ at $\psi=\pi / 2$. This should be zero for any set of $V_{2 m}$ coefficients. By Euler's equation, this condition is obtained if $\partial P / \partial \psi=0$ at this surface. The only quantity that depends on $\psi$ in the preceding is $P_{2 m}(\cos \psi)$. By virtue of being an even function relative to $\psi=\pi / 2$, it follows that $(\partial / \partial \psi) P_{2 m}(\cos \psi) \equiv 0$ at $\psi=\pi / 2$, which verifies that the boundary condition is satisfied identically.

Similar reasoning applies in the case where the hemisphere is mounted on a pressure-release baffle. The pressure in this case is an odd function of $\psi$ relative to $\psi=\pi / 2$, so that $P(r, \pi-\psi)=-P(r, \psi)$. The Legendre functions that fit this symmetry specification are those whose degree is odd. Thus, the field radiated by a hemisphere in an infinite pressure-release baffle is given by

$$
\begin{gather*}
P=\rho_{0} c^{2} \sum_{m=0}^{\infty} V_{2 m+1} \frac{h_{2 m+1}(k r)}{i h_{2 m+1}^{\prime}(k a)} P_{2 m+1}(\cos \psi)  \tag{7.2.29}\\
V_{r}=c \sum_{m=0}^{\infty} V_{2 m+1} \frac{h_{2 m+1}^{\prime}(k r)}{h_{2 m+1}^{\prime}(k a)} P_{2 m+1}(\cos \psi)
\end{gather*}
$$

The odd-degree Legendre functions vanish at the midpoint of their range, $P_{2 m+1}(0) \equiv$ 0 , so the above representation identically satisfies the boundary condition that $P=0$ on the baffle.

The velocity coefficients in Eqs. (7.2.28) and (7.2.29) describe the composite $V(\psi)$ function in Eq. (7.2.26) or (7.2.27). An alternative to integrating over this sphere is to use orthogonality over the hemisphere. The even-degree Legendre polynomials are a mutually orthogonal set over $0 \leq \psi \leq \pi / 2$, as are the odd-degree polynomials. (The even and odd sets are not orthogonal with respect to each other in this range.) Thus, we have

$$
\begin{align*}
\int_{0}^{1} P_{2 m}(\eta) P_{2 j}(\eta) d \eta & =\int_{0}^{\pi / 2} P_{2 m}(\cos \psi) P_{2 j+1}(\cos \psi) \sin \psi d \psi \\
& =\frac{1}{2(2 m)+1} \delta_{j n} \\
\int_{0}^{1} P_{2 m+1}(\eta) P_{2 j+1}(\eta) d \eta & =\int_{0}^{\pi / 2} P_{2 m+1}(\cos \psi) P_{2 j+1}(\cos \psi) \sin \psi d \psi \\
& =\frac{1}{2(2 m+1)+1} \delta_{j n} \tag{7.2.30}
\end{align*}
$$

When we apply these orthogonality properties to the respective series for $V_{r}$, we find that the velocity coefficients are given by

$$
\text { Rigid baffle: } V_{2 m}=(4 m+1) \int_{0}^{\pi / 2} \frac{V(\psi)}{c} P_{2 m}(\cos \psi) \sin \psi d \psi
$$

Pressure-release baffle: $V_{2 m+1}=(4 m+3) \int_{0}^{\pi / 2} \frac{V(\psi)}{c} P_{2 m+1}(\cos \psi) \sin \psi d \psi$
The representations in Eqs. (7.2.28) and (7.2.29) are like those for a sphere in an unbounded space, except that only the even or odd spherical harmonics are relevant. The fact that the domain is a half-space affects the radiated power. The power is obtained by integrating the time-averaged intensity over a large hemisphere. Because $\left(I_{r}\right)_{\mathrm{av}}$ is formed by multiplying $P$ and $V_{r}^{*}$ product, and both are either even or odd
functions relative to $\psi=\pi / 2,\left(I_{r}\right)_{\mathrm{av}}$ is an even function. Thus, the integrals over the hemisphere are half the value of what they would be over the full sphere. Correspondingly, the time-averaged radiated power is obtained by halving Eq.(7.2.17) and using only the appropriate set of $V_{j}$ coefficient. In other words

$$
\mathcal{P}=\left\{\begin{array}{l}
\pi \rho_{0} a^{2} c v_{0}^{2} \frac{1}{(k a)^{2}} \sum_{m=0}^{\infty} \frac{1}{4 m+1} \frac{\left|V_{2 m}\right|^{2}}{\left|h_{2 m}^{\prime}(k a)\right|^{2}}: \text { rigid }  \tag{7.2.32}\\
\pi \rho_{0} a^{2} c v_{0}^{2} \frac{1}{(k a)^{2}} \sum_{m=0}^{\infty} \frac{1}{4 m+3} \frac{\left|V_{2 m+1}\right|^{2}}{\left|h_{2 m+1}^{\prime}(k a)\right|^{2}}: \text { pressure-release }
\end{array}\right.
$$

Somewhere in the midst of this discussion, you might have wondered about the relevance of the concept of an infinite baffle, given that nothing has infinite extent. An important general aspect is that the pressure on a baffle surrounding a radiator tends to fall off rapidly with increasing distance from the center. Thus, the baffle only needs to extend sufficiently far to attain this decrease. In most cases, especially at high frequencies, this distance has the order of magnitude of an acoustic wavelength, which corresponds to an outer radius of the baffle being $R_{\text {baffle }}=a(1+2 \pi / k a)$. In some designs, the apparatus used to mount the transducer is sufficiently large that it is a good approximation of an infinite baffle.

EXAMPLE 7.4 A tweeter consists of a hemisphere that translates as a rigid body at $v_{0} \sin (\omega t)$ in the direction of the axis of symmetry. It is desired to compare the field that this transducer radiates when it is mounted on an infinite rigid or pressure-release baffle to a reference configuration. The latter places the sphere in an unbounded fluid with translational motion only over the forward half, $0 \leq \psi \leq \pi / 2$, with the back half, $\pi / 2<\psi \leq \pi$, stationary. Compare the farfield distribution $(r / a)|p| /\left(\rho_{0} c v_{0}\right)$ for both types of baffles to that obtained from the reference sphere. Also, evaluate the radiated power and the pressure distribution acting on the surface of the hemisphere for each system. Frequencies of interest are $k a=1,4$ and 20.

## Significance

In addition to gaining experience working with spherical harmonics, this example will exhibit some important general features of acoustic radiation and provide insight to the behavior of a common device.

## Solution

In each configuration, the complex amplitude of the radial velocity for $0 \leq \psi \leq \pi / 2$ is $V_{r}=v_{0} \cos \psi$. Both baffle cases give rise to an image of the hemisphere that covers $\pi / 2<\psi \leq \pi$, thereby creating a virtual complete sphere. For a rigid baffle, the radial velocity must be the same on the image, so that $V_{r}=-v_{0} \cos \psi$. The union of these for the virtual sphere is $V_{r}=v_{0}|\cos \psi|$. For a pressure-release baffle, the radial
velocity for $\psi=\pi / 2+\Delta>\pi / 2$ must be equal in magnitude but in the opposite sense from the value at $\psi=\pi / 2-\Delta$. Thus, the radial velocity on the virtual sphere is $V_{r}=v_{0} \cos \psi$. The radial velocity on the reference full sphere is $V_{r}=v_{0} \cos \psi$ for $0 \leq \psi \leq \pi / 2$ and $V_{r}=0$ for $\pi / 2<\psi \leq \pi$.

The radial velocity on the virtual sphere for the rigid baffle case is an even function with respect to $\psi=\pi / 2$. This means that only the only nonzero velocity coefficients are $V_{2 m}, m=0,1,2, \ldots$. According to the first of Eq. (7.2.31), these coefficients are

$$
V_{2 m}=(4 m+1) \int_{0}^{\pi / 2}(\cos \psi) P_{2 m}(\cos \psi) \sin \psi d \psi
$$

This is like Eq. (1) in Example 7.3 with $\psi_{0}=\pi / 2$, except that $2 m$ replaces $m$ and the integral is doubled. When we make these alterations, the result is

$$
\begin{aligned}
\left(V_{2 m}\right)_{\mathrm{rigid}}= & -\left[\frac{2 m+1}{4 m+3} P_{2 m+2}(0)+\frac{(4 m+1)}{(4 m+3)(4 m-1)} P_{2 m}(0)\right. \\
& \left.-\frac{2 m}{4 m-1} P_{2 m-2}(0)\right], m>0 \\
\left(V_{0}\right)_{\text {rigid }}= & \frac{1}{2}
\end{aligned}
$$

In the case of a pressure-release baffle, the radial velocity function for the equivalent full sphere is an odd function relative to $\psi=\pi / 2$. In other words, it is $\cos \psi$. Thus, the surface velocity of the image sphere is that of a sphere that translates as a rigid body. Because $P_{1}(\cos \psi) \equiv \cos \psi$, orthogonality of the Legendre functions informs us that $V_{1}=1$ is the only nonzero coefficient.

The radial velocity on the reference full sphere has no special symmetry properties with respect to $\psi=\pi / 2$. The velocity coefficients are given directly by Eq.(7.2.4) for $m=0,1,2, \ldots$ Setting $V_{r}=0$ for $\psi>\pi / 2$ leads to

$$
V_{m}=\left(m+\frac{1}{2}\right) \int_{0}^{1} \eta P_{m}(\eta) d \eta
$$

This equation is the same as Eq. (1) in Example 7.3 with $\psi_{0}=\pi / 2$, so we quote the results directly as

$$
\begin{aligned}
\left(V_{m}\right)_{\mathrm{ref}}= & -\frac{1}{2} \frac{v_{0}}{c}\left[\frac{m+1}{2 m+3} P_{m+2}(0)+\frac{(2 m+1)}{(2 m+3)(2 m-1)} P_{m}(0)\right. \\
& \left.-\frac{m}{2 m-1} P_{m-2}(0)\right], m>1 \\
\left(V_{0}\right)_{\mathrm{ref}}= & \frac{1}{4} \frac{v_{0}}{c}, \quad V_{1}=\frac{1}{2} \frac{v_{0}}{c}
\end{aligned}
$$

Despite the similarity of the surface velocity in the three configurations, the $V_{j}$ coefficients are significantly different, as is shown in Fig. 1. This figure also shows that truncation of the spherical harmonic series at $N=12$ should be adequate for the rigid and reference cases.


Figure 1.

The computation of the directivity proceeds as though we had a full sphere, except that $\psi$ only extends to $\pi / 2$ when there is a baffle, and the velocity coefficients are those of the relevant configuration. Equation (7.2.14) for the three cases gives

$$
\begin{gathered}
\mathcal{D}_{\text {rigid }}(\psi)=\frac{1}{k a}\left|\sum_{m=0}^{\infty} \frac{\left(V_{2 m}\right)_{\text {rigid }}}{i h_{2 m}^{\prime}(k a)} e^{i(2 m+1) \pi / 2} P_{2 m}(\cos \psi)\right| \\
\mathcal{D}_{\text {p.r. }}(\psi)=\frac{1}{k a}\left|\frac{\left(V_{1}\right)_{\text {p.r. }}}{i h_{1}^{\prime}(k a)}\right| \cos \psi=\frac{v_{0}}{c} \frac{(k a)^{2}}{\left|2+2 i k a-(k a)^{2}\right|} \cos \psi \\
\mathcal{D}_{\text {ref }}(\psi)=\frac{1}{k a}\left|\sum_{m=0}^{\infty} \frac{\left(V_{m}\right)_{\text {ref }}}{i h_{m}^{\prime}(k a)} e^{i(m+1) \pi / 2} P_{m}(\cos \psi)\right|
\end{gathered}
$$

The directivities for each specified frequency are displayed in Fig. 2.


Figure 2.

Because the pressure-release baffle corresponds to an image sphere that oscillates as a rigid body in the $z$ direction, the directivity is always proportional to $\cos \psi$, but its maximum value is frequency dependent. At the lowest frequency, $k a=1$, the rigid baffle and the reference sphere radiate approximately uniform. This is consistent with the general observation that a body that executes a low-frequency vibration acts like a monopole whose strength is proportional to the body's volume velocity. In contrast, at the highest frequency, $k a=20$, each directivity fits the asymptotic trend that it will be proportional to the surface velocity distribution.

The radiated power in each case may be found from Eq. (7.2.17) using the appropriate set of $V_{j}$ coefficients. For the configurations where there is a baffle, the
formula is halved because a baffle reduces the domain to a half-space. The results for $\mathcal{P} /\left(\pi \rho_{0} a^{2} c v_{0}^{2}\right)$ are tabulated below.

|  | $\mathcal{P}$ (nondim.) |  |  |
| :---: | :---: | :---: | :---: |
|  | 1 | 4 | 20 |
| Rigid baffle | 0.12588 | 0.32242 | 0.33331 |
| Pressure-release baffle | 0.066667 | 0.32821 | 0.33333 |
| Reference sphere | 0.096272 | 0.32531 | 0.33332 |

At low frequencies, the rigid baffle configuration gives twice as much power radiation as the pressure-release case, while the reference sphere is midway between the two baffle configurations. In contrast, at high frequencies, all configurations give the same radiated power. This is to be expected because the directivity of all configurations at large $k a$ is essentially the same for $\psi<90^{\circ}$, and it is close to zero for $\psi>90^{\circ}$ in the case of the reference sphere.

The evaluation of surface pressure follows a direct implementation of Eq. (7.2.7), with the $V_{j}$ coefficients for the rigid baffle case being those for even $j$, while the only nonzero coefficient for the pressure-release baffle is $V_{1}$. Of course, only $0 \leq \psi \leq \pi / 2$ is relevant. Here too, the matrix algorithm for the evaluation of a series representation may be applied.

The radial velocity on the surface is real and positive. Thus, $\operatorname{Re}(P)$ is the resistive part, and $\operatorname{Im}(P)$ is the reactive part, with $\operatorname{Im}(P)>0$ for an inertance. The patterns in Fig. 3 indicate that when the frequency is low, the surface pressure for both types of baffles has a large inertance. The resistive part for the rigid baffle at low frequency is approximately uniform, whereas $\operatorname{Re}(P)$ is proportional to $\cos \psi$ for the pressurerelease baffle. High frequencies lead to a surface pressure that is $\rho_{0} c v_{0} \cos \psi$ for both configurations. This occurs because the surface, $r=a$, effectively is in the farfield when $k a$ is very large.


In systems for sound reproduction, it usually is desirable to emit the sound omnidirectionally. This is the role of a hemispherical dome tweeter. Because $a$ is relatively small (typically less than 40 mm ), the value of $k a$ will be small even at the upper limits of the audible spectrum. Thus, their use leads to radiation of high frequencies over a broad angular range. The rigid baffle enhances the radiated power at low frequencies. A designer would be interested in the surface pressure for design purposes. For example, the fact there is a large reactive part means that the electronic amplifier that drives the tweeter must be able to output an instantaneous power that is substantially greater than the time-averaged power.

### 7.2.4 Interaction with an Elastic Spherical Shell

A situation we have not yet addressed arises when the surface vibration is produced by a force applied to an elastic body. The vibration in that case must be determined by solving laws of structural dynamics. Such an analysis must account for the pressure applied to the surface by the acoustic signal that is generated. But that pressure depends on the surface vibration. This is not circular logic. Rather, it requires a simultaneous approach that concurrently formulates the governing acoustic and structural equations. Problems such as this fall into the category of fluid-structure interaction. If the shape of the body and the nature of the structure are complicated, as in the case of a submarine, then accurate determination of the radiated field at all frequencies would require the most powerful computational tools available, and even such resources might not be adequate. The decoupling properties of spherical harmonics make the analysis of radiation from elastic spheres quite tractable.

We previously encountered an elastic plate, which refers to a planar sheet whose thickness is much less than the dimensions of the sheet. A shell consists of a thin sheet whose surface is curved. Doing so alters the nature of the internal stresses in comparison to a plate, which thereby complicates the equations of motion. We will begin with a membrane shell model, in which the internal stresses come from stretching the sheet.

To some extent, the stresses resemble how a balloon carries a load, but there is a fundamental difference. In the case of a balloon, the sheet is inflated before it is subject to loading. Inflation stretches the sheet, thereby inducing a prestress. Displacement of the sheet produces additional stress. The work that is done in this displacement, which is stored as potential energy, comes from displacing the system in opposition to the total stress. In the case of a membrane shell, there is no prestress, so the internal work stems from displacing the shell in opposition to the stretching stress induced by that displacement. A simpler system that has the same alternatives is a cable. The inflated balloon is analogous to a tensioned cable, whereas a membrane shell is like a taut, but untensioned cable. Interestingly, the taut cable has no load carrying ability according to linear elasticity theory. This is not true for the membrane shell because of its curvature.

The membrane shell theory we will employ becomes progressively less valid as the thickness/curvature ratio increases. It also is not valid if the frequency is so high that the wavelength of the surface displacement is comparable to the thickness. Improved shell theories that address both situations are available. (The author once joked to a colleague that there might be more improved shell theories than the number of people who have been involved in the development of those theories.) We shall begin with membrane shell theory because it is the simplest model. However, a breakdown of this theory will cause an anomaly of the series, which we will address by introducing a correction that accounts for flexural effects.

## Equations of Motion

Figure 7.4 shows the midsurface of a spherical shell, which is the surface that is midway between the outer and inner surfaces. We will limit our consideration to situations in which the shell response is axisymmetric. This means that the displacement of a point at $(a, \psi)$ on the midsurface will consist of a radial component $w$ and meridional displacement $u$ tangent to the surface. A positive displacement component is one that moves the midsurface in the sense of $\bar{e}_{r}$ or $\bar{e}_{\psi}$. The restriction to axisymmetric motions limits $u$ and $w$ to be functions of $\psi$ and $t$ only. The shell's thickness $h$ is much less than $a$. Thus, we may also use $a$ as the radius of the outer surface when we formulate the equations governing the acoustic field exterior to the sphere.

Fig. 7.4 Definition of the radial displacement $w$ and polar displacement $u$ of a spherical shell


Junger and Feit ${ }^{7}$ employed equations of motion that incorporate the effects of flexural rigidity. Omitting those effects leads to the following simplified set of equations of membrane shell theory,

[^5]\[

$$
\begin{gather*}
c_{\mathrm{e}}^{2}\left[\frac{\partial^{2} u}{\partial \psi^{2}}+\cot \psi \frac{\partial u}{\partial \psi}-\left(\nu+(\cot \psi)^{2}\right) u+(1+\nu) \frac{\partial w}{\partial \psi}\right]-a^{2} \frac{\partial^{2} u}{\partial t^{2}}=0 \\
-c_{\mathrm{e}}^{2}(1+\nu)\left[\frac{\partial u}{\partial \psi}+(\cot \psi) u+2 w\right]-a^{2} \frac{\partial^{2} w}{\partial t^{2}}=\frac{a^{2}}{\rho_{\mathrm{e}} h}\left(\left.p_{\text {acoustic }}\right|_{r=a}-q_{\text {applied }}\right) \tag{7.2.33}
\end{gather*}
$$
\]

In this expression, $q_{\text {applied }}$ represents the outward radial force per unit surface area that is the excitation, $\rho_{\mathrm{e}}$ is the density of the shell's material, $\nu$ is Poisson's ratio, and $c_{\mathrm{e}}$ is the phase speed of an extensional wave of plane strain in a plate of the same thickness,

$$
\begin{equation*}
c_{\mathrm{e}}=\left[\frac{E}{\rho_{\mathrm{e}}\left(1-\nu^{2}\right)}\right]^{1 / 2} \tag{7.2.34}
\end{equation*}
$$

We know that the pressure exterior to a sphere may be represented by a spherical harmonic series. Also, the acceleration term in the second of Eq. (7.2.33) is proportional to $w$ in the case of a harmonic response. Because the $w$ term in that equation is not differentiated with respect to $\psi$, it seems reasonable to try a Legendre function series for $w$. In both equations of motion, most of the $u$ terms differ from the $w$ terms by one differentiation with respect to $\psi$. This feature suggests that we should try a series representation of $u$, with the derivatives of the Legendre functions as the basis functions. Thus, the ansatz we begin with is

$$
\begin{gather*}
p=\operatorname{Re} \sum_{m=0}^{\infty} B_{m} h_{m}(k r) P_{m}(\cos \psi) e^{i \omega t} \\
w=\operatorname{Re} \sum_{m=0}^{\infty} W_{m} P_{m}(\cos \psi) e^{i \omega t}  \tag{7.2.35}\\
u=\operatorname{Re} \sum_{m=0}^{\infty} U_{m} \frac{d}{d \psi} P_{m}(\cos \psi) e^{i \omega t}
\end{gather*}
$$

Many mathematical manipulations are required to prove that the basis functions for $w$ and $u$ are consistent with the equations of motion, and to identify the equations governing the coefficients $W_{m}$ and $U_{m}$. Both Baker ${ }^{8}$ and Junger and Feit ${ }^{9}$ apply the transformation $\eta=\cos \psi$ prior to identifying the basis functions. In both analyses, the series for $w$ uses $P_{m}(\eta)$, which is equivalent to the above. The $u$ basis function used by Junger and Feit is $\left(1-\eta^{2}\right)^{1 / 2} P_{m}^{\prime}(\eta)$, whereas the basis function for $u$ in Baker's analysis is the associated Legendre polynomial $P_{m}^{1}(\eta)$. Equation (7.1.42) states this function is $P_{m}^{1}(\eta)=(d / d \psi) P_{m}(\cos \psi)=-\left(1-\eta^{2}\right)^{1 / 2} P_{m}^{\prime}(\eta)$. Thus, Baker's series matches the one used here, whereas Junger and Feit's work corresponds to $U_{m}$ coefficients that are the negative of the present definition.

[^6]The equations that result from substitution of Eq. (7.2.35) into the equations of motion, followed by application of the derivative identities in Eq. (7.1.14), are

$$
\begin{align*}
& \sum_{m=0}^{\infty}\left\{\left[-K_{U U}+\left(k_{\mathrm{e}} a\right)^{2}\right] U_{m}+K_{U W} W_{m}\right\} P_{m}^{1}(\eta)=0 \\
& \sum_{m=0}^{\infty}\left\{K_{W U} U_{m}+\left[-K_{W W}+\left(k_{\mathrm{e}} a\right)^{2}\right] W_{m}\right\} P_{m}(\eta) \\
& =\frac{a^{2}}{\rho_{s} h c_{\mathrm{e}}^{2}}\left[\sum_{m=0}^{\infty} B_{m} h_{m}(k a) P_{m}(\eta)-\hat{q}_{\text {applied }}\right] \tag{7.2.36}
\end{align*}
$$

where the $K$ coefficients are functions of $m$ given by

$$
\begin{align*}
& K_{U U}=m(m+1)-(1-\nu), \quad K_{U W}=(1+\nu) \\
& K_{W U}=(1+\nu) m(m+1), \quad K_{W W}=2(1+\nu) \tag{7.2.37}
\end{align*}
$$

The wavenumber $k_{\mathrm{e}}$ is defined relative to $c_{\mathrm{e}}$, rather than $c$, in order to avoid confusion when we consider the shell in a vacuum. Thus,

$$
\begin{equation*}
k_{\mathrm{e}}=\frac{\omega}{c_{\mathrm{e}}}, \quad k=\frac{c_{\mathrm{e}}}{c} k_{\mathrm{e}} \tag{7.2.38}
\end{equation*}
$$

The first of Eq. (7.2.36) yields uncoupled equations because the $P_{m}^{1}(\eta)$ are a linearly independent set. To uncouple the second equation, we employ the orthogonality property of Legendre functions, Eq. (7.1.16). The result of multiplying the equation by a specific $P_{n}(\eta)$, then integrating over $-1 \leq \eta \leq 1$, is

$$
\begin{gather*}
{\left[K_{U U}-\left(k_{\mathrm{e}} a\right)^{2}\right] U_{m}-K_{U W} W_{m}=0} \\
-K_{W U} U_{m}+\left[K_{W W}-\left(k_{\mathrm{e}} a\right)^{2}\right] W_{m}=-\frac{a^{2}}{\rho_{e} h c_{\mathrm{e}}^{2}} B_{m} h_{m}(k a)+\frac{a^{2}}{\rho_{s} h c_{\mathrm{e}}^{2}} F_{m} \tag{7.2.39}
\end{gather*}
$$

The $F_{m}$ coefficients are those of a Legendre series for the applied excitation,

$$
\begin{equation*}
F_{m}=\frac{2 m+1}{2} \int_{-1}^{1} \hat{q}_{\text {applied }} P_{m}(\eta) d \eta \tag{7.2.40}
\end{equation*}
$$

These values are considered to be known, so there are three unknowns in Eq. (7.2.39).
The additional equation comes from enforcing continuity at the sphere's surface. The relation between the spherical harmonic coefficients of pressure and radial velocity on a sphere's surface is provided by Eq. (7.2.6). The radial velocity of the surface is $\dot{w}$, so the velocity coefficients are $i \omega W_{m}$, from which it follows that

$$
\begin{equation*}
B_{m}=\rho_{0} c \omega \frac{1}{h_{m}^{\prime}(k a)} W_{m} \tag{7.2.41}
\end{equation*}
$$

We now have a solvable triad of equations. The most direct solution uses the preceding to eliminate $B_{m}$ and the first of Eq. (7.2.36) to eliminate $U_{m}$. This leads to

$$
\begin{gather*}
U_{m}=\frac{K_{U W}}{K_{U U}-\left(k_{\mathrm{e}} a\right)^{2}} W_{m} \\
{\left[K_{W W}-\left(k_{\mathrm{e}} a\right)^{2}-\frac{K_{U W} K_{W U}}{K_{U U}-\left(k_{\mathrm{e}} a\right)^{2}}+\frac{\rho_{0} a}{\rho_{s} h}\left(\frac{c}{c_{\mathrm{e}}}\right)^{2} \frac{k a h_{m}(k a)}{h_{m}^{\prime}(k a)}\right] \frac{W_{m}}{a}=\frac{a}{h} \frac{F_{m}}{\rho_{\mathrm{e}} c_{\mathrm{e}}^{2}}} \tag{7.2.42}
\end{gather*}
$$

Terms have been grouped to emphasize the manner in which the system parameters affect the response. The pressure excitation is referenced to $\rho_{\mathrm{e}} c_{\mathrm{e}}^{2}$, which is a bulk modulus for the shell, and displacement is referenced to the radius. This nondimensional displacement depends parametrically on the density and sound speed ratios, Poisson's ratio $\nu$, and the nondimensional frequency.

Evaluation of the response at any frequency entails solving the second of the above equations for the radial displacement coefficient $W_{m}$ for a range of harmonic numbers $m$, then using those coefficients to evaluate the corresponding pressure coefficients $B_{m}$ according to Eq. (7.2.41). The meridional displacement typically is not of interest, but if it were, we would use the first equation above to evaluate the $U_{m}$ values. The displacement and pressure fields would be synthesized according to Eq. (7.2.35). The farfield description of the pressure is obtained by replacing $h_{m}(k r)$ by its asymptotic behavior for large $k r$, which yields

$$
\begin{equation*}
p=\operatorname{Re}\left\{\frac{e^{-i k r}}{k r} \sum_{m=0}^{\infty} B_{m} e^{i(m+1) \pi / 2} P_{m}(\cos \psi) e^{i \omega t}\right\} \tag{7.2.43}
\end{equation*}
$$

## In-vacuo Vibration

An important property of the shell is its natural frequencies, at which the system may vibrate without application of an external excitation. No free system can execute a steady-state harmonic response if energy is removed from it. As we will soon see, that is what the fluid always does. Thus, the natural frequencies correspond to the harmonic motion of the in-vacuo shell when there is no excitation. The effect of fluid loading is described by the last term in the bracket in the second of Eq. (7.2.42). Dropping this term and setting $F_{m}=0$ leads to the characteristic equation, which is a quadratic equation whose roots are the natural frequencies $\Omega_{m}$ of the $m$ th spherical harmonic,


Fig. 7.5 Natural frequencies and modal displacement ratios of a spherical shell according to membrane shell theory

$$
\begin{equation*}
\chi\left(\frac{\Omega_{m} a}{c_{\mathrm{e}}}\right)=\left[K_{W W}-\left(\frac{\Omega_{m} a}{c_{\mathrm{e}}}\right)^{2}\right]\left[K_{U U}-\left(\frac{\Omega_{m} a}{c_{\mathrm{e}}}\right)^{2}\right]-K_{U W} K_{W U}=0 \tag{7.2.44}
\end{equation*}
$$

Thus, there are two natural frequencies $\Omega_{m}^{(1)}<\Omega_{m}^{(2)}$ at a specific $m$. At either natural frequency, Eq. (7.2.39) with $B_{m}=F_{m}=0$ ceases to be linearly independent. Consequently, we may only determine the ratio $U_{m} / W_{m}$ from either of those equations. When the natural frequencies are plotted as functions of $m$, they form two curves called branches. They are depicted in Fig. 7.5 along with the displacement ratio $U_{m} / W_{m}$ for each branch. The common parlance is to say that the displacement pattern in a free vibration at a system's natural frequency is a mode function or mode shape. The radial displacement in a mode is the product of $W_{m}$ and $P_{m}(\cos \psi)$, while the meridional displacement is the product of $U_{m}$ and $(d / d \psi) P_{m}(\cos \psi)$.

The lower branch has no natural frequency at $m=0$. (The mathematical solution is not physical, because it corresponds to an imaginary natural frequency.) At $m=1$, the lower natural frequency is zero, which means that it corresponds to a rigid body mode in which the shell moves without deformation. This is verified by the modal displacement ratio $U_{1}^{(1)} / W_{1}^{(1)}=1$, which corresponds to $w=\operatorname{Re}\left(W_{1}^{(1)} \cos \psi e^{i \Omega^{(1)} t}\right), u=\operatorname{Re}\left(-W_{1}^{(1)} \sin \psi e^{i \Omega^{(1)} t}\right)$. This pattern is translational oscillation in the axial direction.

If the natural frequencies are known, the second of Eq.(7.2.42) in the case of forced response may be written as
where

$$
\begin{equation*}
\kappa_{m}^{(1)} \equiv \Omega_{m}^{(1)} \frac{a}{c}, \quad \kappa_{m}^{(2)} \equiv \Omega_{m}^{(2)} \frac{a}{c}, \quad \kappa_{m}^{(0)}=\left(K_{U U}\right)^{1 / 2} \tag{7.2.46}
\end{equation*}
$$

The special case of a shell in a vacuum is obtained by setting $\rho_{0} / \rho_{e}=0$. It is evident that removal of fluid loading will cause $W_{m}$ to be singular when $k_{\mathrm{e}} a$ equals $\kappa_{m}^{(1)}$ or $\kappa_{m}^{(2)}$. This condition marks a resonance, which we will examine more closely in the following section. In contrast, this equation indicates that $W_{m}$ will be zero when $k_{\mathrm{e}} a=\kappa_{m}^{(0)}$. In control system theory, $\omega=\Omega_{m}^{(1)}$ or $\omega=\Omega_{m}^{(2)}$ are "poles" of the in-vacuo system, and $\omega=\left(c_{\mathrm{e}} / a\right)\left(K_{U U}\right)^{1 / 2}$ is a "zero".

A troubling aspect of the displacement in any case becomes evident if we consider the value of $W_{m}$ as $m$ is increased with $k_{\mathrm{e}} a$ fixed. According to Fig. 7.5, $\kappa_{m}^{(1)}$ tends to a constant value as $m$ increases. Also, examination of Eq. (7.2.37) indicates that $K_{U U}$ approaches $m^{2}$, so $\kappa_{m}^{(0)}$ approaches $m$. Consequently, for very large $m$ at fixed $k a$, the first term in Eq. (7.2.45) approaches a constant value. The asymptotic expansion of $h_{m}(k a)$ and its derivative for large $m$, see Eq.(7.1.36), indicate that the second term also tends to a constant value. If the $F_{m}$ coefficients do not decrease with increasing $m$, the $W_{m}$ values will not decrease. In that event, none of the spherical harmonic series will converge.

The source of this difficulty is the usage of membrane shell theory. In particular, the fault lies in its prediction that the natural frequencies on the lower branch $\kappa_{m}^{(1)}$ approach a constant value as $m$ is increased. Any shell theory that incorporates flexural deformation effects will yield natural frequencies that increase with increasing harmonic number. For this reason, we shall henceforth use the full theory presented by Junger and Feit. The parameter that scales flexural effects relative to membrane effects is

$$
\begin{equation*}
\beta^{2}=\frac{h^{2}}{12 a^{2}} \tag{7.2.47}
\end{equation*}
$$

Fortunately, adoption of this improved shell theory only entails altering the definitions of the $K$ coefficients, which now are

$$
\begin{gather*}
K_{U U}=\left(1+\beta^{2}\right)[m(m+1)-(1-\nu)] \\
K_{U W}=(1+\nu)+\beta^{2}[m(m+1)-(1-\nu)] \\
K_{W U}=m(m+1)\left[(1+\nu)+\beta^{2} m(m+1)-\beta^{2}(1-\nu)\right]  \tag{7.2.48}\\
K_{W W}=2(1+\nu)+\beta^{2} m^{2}(m+1)^{2}-\beta^{2}(1-\nu) m(m+1)
\end{gather*}
$$



Fig. 7.6 Comparison of natural frequencies of a spherical shell according to membrane theory to those obtained when flexural effects are included, $a / h=50$

The natural frequencies that result from inclusion of flexural effects are depicted in Fig. 7.6. There is no perceptible effect on the natural frequencies of the upper branch. The frequencies of the lower branch increase as expected, and the rate of increase is more rapid than a simple proportionality to $m$. This increase is sufficient to assure that the $W_{m}$ values obtained from Eq. (7.2.45) decrease for sufficiently large $m$, regardless of the nature of the excitation, $q_{\text {applied }}$. Decreasing $a / h$ increases the values of $\Omega_{m}^{(1)}$, but has little effect on the values of $\Omega_{m}^{(2)}$. The modal displacement ratios $U_{m}^{(j)} / W_{m}^{(j)}$ are not displayed because the values associated with either branch are indistinguishable from those in Fig. 7.5.

Even with the inclusion of flexural effects, an important limitation must be recognized. Simple shell theory is not valid if the wavelength along the surface is comparable to the thickness, because other deformation effects must be included. A qualitative guideline for validity may be identified from the property that $P_{m}(\cos \psi)$ has $m$ zeros in the range $0<\psi<\pi$. If we ignore the fact that these zeros are not equidistant, we arrive at a surface half-wavelength of $2 \pi a / \mathrm{m}$. A reasonable criterion is that this distance should be no less than $10 h$. This leads to a guideline that the shell theory employed here can be expected to lose accuracy if $m>(\pi / 5) a / h$. For example, if $a / h=50$, it is desirable that the spherical harmonic series converge prior to $m=32$. If they do not, we might obtain a mathematically convergent solution that does not accurately describe the actual physical response.

## Interpretation

In the present context, the in-vacuo vibration properties provide a lens for interpreting the fluid-loaded response. To see why, let us compare Eq. (7.2.42) to the expression for the displacement of a one-degree-of-freedom oscillator subjected to a harmonically varying force. Let $K, M$, and $D$, respectively, denote stiffness, mass, and dashpot constants, nondimensionalized based on the usage of $k a$ to represent frequency. In terms of a nondimensional velocity $V=i k a(W / a)$, the steady-state amplitude of the oscillator is governed by

$$
\begin{equation*}
Z V=F, \quad Z=i\left(k a M-\frac{K}{k a}\right)+D \tag{7.2.49}
\end{equation*}
$$

The imaginary part of the impedance is the reactance. It is negative if $k a$ is below the natural frequency, which is $(K / M)^{1 / 2}$, and it is positive if $k a$ exceeds the natural frequency. The dashpot term is a positive real value at any frequency. If the dashpot was not present, a frequency sweep would reveal a true resonance at the natural frequency, in which the value of $V$ is infinite. ${ }^{10}$ The dashpot removes energy at any frequency, so the response never shows a true resonance if the dashpot is present.

To apply this perspective to the spherical shell, we convert Eq. (7.2.45) to an impedance form by factoring $i k_{\mathrm{e}} a$ out of the coefficient. The coefficients in the first line of this expression represent the effects of the structure, whereas the term in the second line is due to the fluid. Let us denote these contributions as the structural impedance $\left(Z_{\mathrm{e}}\right)_{m}$ and the fluid surface impedance $\left(Z_{\mathrm{f}}\right)_{m}$, respectively. Thus, we rewrite the equation as

$$
\begin{gather*}
{\left[\left(Z_{\mathrm{e}}\right)_{m}+\left(Z_{\mathrm{f}}\right)_{m}\right]\left(i k_{\mathrm{e}} a \frac{W_{m}}{a}\right)=\frac{a}{h} \frac{F_{m}}{\rho_{\mathrm{e}} c_{\mathrm{e}}^{2}}}  \tag{7.2.50}\\
\left(Z_{\mathrm{e}}\right)_{m}=\frac{k_{\mathrm{e}} a}{i} \frac{\left[\left(\kappa_{m}^{(1)}\right)^{2} /\left(k_{\mathrm{e}} a\right)^{2}-1\right]\left[\left(\kappa_{m}^{(2)}\right)^{2} /\left(k_{\mathrm{e}} a\right)^{2}-1\right]}{\left(\kappa_{m}^{(0)}\right)^{2} /\left(k_{\mathrm{e}} a\right)^{2}-1} \\
\left(Z_{\mathrm{f}}\right)_{m}=-i \frac{\rho_{0} a}{\rho_{e} h}\left(\frac{c}{c_{e}}\right) \frac{h_{m}(k a)}{h_{m}^{\prime}(k a)}
\end{gather*}
$$

The frequency dependence of $\left(Z_{\mathrm{e}}\right)_{m}+\left(Z_{\mathrm{f}}\right)_{m}$ does not match that of a simple oscillator, but we can see that $\left(Z_{\mathrm{e}}\right)_{m}$ is imaginary, so it is a reactance. If $k_{\mathrm{e}} a<\kappa_{m}^{(1)}$, then it is negative imaginary, corresponding to a compliance. Also, $\left(Z_{\mathrm{e}}\right)_{m}$ is infinite when $k_{e} a=\kappa_{m}^{(0)}$, and it grows without bound to a large positive imaginary value (inertance) with increasing $k_{\mathrm{e}} a$ beyond $\kappa_{m}^{(2)}$. A large impedance leads to small velocity and displacement amplitudes.

The fluid impedance, $\left(Z_{\mathrm{f}}\right)_{m}$, is complex and nonsingular. The frequency dependence of both impedances for the lowest spherical harmonics is described by Fig. 7.7 for the case of a steel spherical shell submerged in water. Several features are noteworthy. The most important are that $\operatorname{Im}\left(Z_{f}\right)>0$ and $\operatorname{Re}\left(Z_{f}\right)>0$ for any frequency. In other words the fluid loading adds an inertance and a resistance.

We have seen that if the shell vibrates in a vacuum, it will resonate when $k_{\mathrm{e}} a=\kappa_{m}^{(1)}$ or $k_{\mathrm{e}} a=\kappa_{m}^{(2)}$. Both conditions correspond to $\left(Z_{\mathrm{e}}\right)_{m}=0$. (In an actual system, internal dissipation would limit $W_{m}$ to a finite value.) Now consider what happens when the impedance of the fluid is added to that of the shell. Because $\operatorname{Im}\left(Z_{f}\right)$ is always greater than zero, a plot of $\operatorname{Im}\left(\left(Z_{\mathrm{e}}\right)_{m}+\left(Z_{\mathrm{f}}\right)_{m}\right)$ as a function of $k_{\mathrm{e}} a$ would yield a curve at each $m$ that is above the corresponding plot of $\operatorname{Im}\left(Z_{\mathrm{e}}\right)_{m}$ in Fig. 7.7. Consequently, the total inertance crosses the zero axis at a lower frequency. It is convenient to refer

[^7]

Fig. 7.7 Structural and fluid impedances for several spherical harmonics as a function of frequency. The media are steel and water, and $a / h=50$
to any condition in which the total reactance vanishes as a fluid-loaded resonance. These resonances occur at a lower frequency than the corresponding (in-vacuo) natural frequencies. Despite the terminology, it is essential that one not forget that fluid resistance is always present. This resistance, being in-phase with the surface velocity, plays the same role as a dashpot. The difference is that a dashpot dissipates energy internally, whereas the power required to overcome the fluid resistance is transferred into the fluid as the radiated pressure field. The name given to this effect is radiation damping.

Another important observation stems from the fact that the magnitude of $Z_{f}$ is proportional $\left(\rho_{0} c\right) /\left(\rho_{\mathrm{e}} c_{\mathrm{e}}\right)$, that is, the ratio of characteristic impedances. For water and steel, this ratio is approximately 0.03 , whereas it is approximately $0.9\left(10^{-5}\right)$ in the case of air and steel. The usual terminology is to say that the former constitutes heavy fluid loading, whereas the latter is light fluid loading. In a lightly fluid-loaded system, the fluid impedance will be much smaller than the structural impedance, except at the in-vacuo natural frequencies, where $\operatorname{Im}\left(Z_{\mathrm{e}}\right)=0$.

This attribute leads to a simplification that often is invoked in the cases of light fluid loading. In it, the structural response is determined with the in-vacuo model. The effect of fluid resistance when the frequency is close to in-vacuo natural frequencies is estimated, typically on the basis of prior experience with similar systems. The vibratory displacement is then used to predict the pressure field as a standard radiation problem. The important aspect of this simplified approach is that increasing $\rho_{0} c / \rho_{\mathrm{e}} c_{\mathrm{e}}$ will lead to a decrease in the peak amplitudes at resonances, but it will not account for the downshift of the frequency at which these resonances occur.

The discussion thus far has focused on the behavior of individual spherical harmonics. We have examined the factors influencing the displacement and pressure coefficients, and have seen that the structure has an infinite number of in-vacuo modes that have some relevance to resonant-type phenomena. We have not seen how these behaviors combine to affect the displacement and acoustic responses to an actual excitation. This we shall do by investigating a prototypical system.

> EXAMPLE 7.5 A spherical shell $(a / h=40)$ composed of aluminum is submerged in water. It is excited by a time-harmonic point force that acts in the radial direction. (a) Determine the amplitude of the displacement coefficients $W_{m}$ as a function of frequency for $m \leq 4$ and $0<k_{\mathrm{e}} a<8$. (b) Determine the amplitude of the radial displacement at the point where the force is applied and at the diametrically opposite point for $0<k_{\mathrm{e}} a<8$. (c) Determine the farfield pressure amplitude $r|P|$ for $0<k_{\mathrm{e}} a<8$ in the direction $\psi=0$ and $\psi=180^{\circ}$, where $\psi$ is measured relative to the radial line through the point where the force is applied. (d) Determine the directivity of the farfield pressure amplitude. Carry out the analysis for three frequencies: the frequency in the range $k_{\mathrm{e}} a<1$ for which $r|P|$ at $\psi=0$ is maximized, $k_{\mathrm{e}} a=2.5$, and the frequency in the range $k_{\mathrm{e}} a<5$ for which $r|P|$ at $\psi=0$ is maximized.

## Significance

An important part of any analysis is interpretation and explanation of the results. This example will explore how to do so when elastic and acoustic effects interact. Along the way, issues regarding efficient computational approaches and series convergence will come to the fore.

## Solution

The only deviation of this system from radial symmetry is the point force. If its location is defined as the $\psi=0$ pole for a spherical coordinate system, the system is axisymmetric with respect to the $z$-axis. The value of $W_{m}$ as a function of frequency is given by Eq. (7.2.42). Solution of those equations requires specification of the $F_{m}$ coefficients, which are described by Eq. (7.2.40). That expression is based on a distributed force, that is, a force per unit area, whereas the present excitation is a concentrated force $\operatorname{Re}\left(\hat{F} e^{i \omega t}\right)$ applied at $\eta=0$. We could use a Dirac delta to describe this force, but a less mathematical approach is available. Let us consider the force to be distributed over a small spherical cap defined by $0 \leq \psi \leq \Delta$, where $\Delta \ll 1$. The point force coefficients will emerge by recognizing the smallness of $\Delta$.

The distributed force is $q_{\text {applied }}=\hat{F} / \mathcal{A}$ for $\psi \leq \Delta$ and $q_{\text {applied }}=0$ for $\psi>\Delta$, where $\mathcal{A}$ is the area of the cap,

$$
\mathcal{A}=\int_{0}^{\Delta}(2 \pi a \sin \psi)(a d \psi)=2 \pi a^{2}(1-\cos \Delta) \approx \pi a^{2} \Delta^{2}
$$

It is convenient to use $\eta=\cos \psi$ to change the integration variable in Eq. (7.2.40), so that

$$
F_{m}=\frac{2 m+1}{2} \int_{0}^{\Delta} \frac{\hat{F}}{\pi a^{2} \Delta^{2}} P_{m}(\cos \psi) \sin \psi d \psi
$$

Because $\Delta$ is extremely small, the integration variable is always small. Therefore, the Legendre functions are well approximated as $P_{m}(\cos \psi) \approx 1$, and we may set $\sin \psi \approx \psi$. The result is

$$
\begin{equation*}
F_{m}=\left(\frac{2 m+1}{4}\right) \frac{\hat{F}}{\pi a^{2}} \tag{1}
\end{equation*}
$$

Before we proceed to any computations, it is useful to review the equations to be solved. First, the nondimensional natural frequencies $\kappa_{m}^{(1)}$ and $\kappa_{m}^{(2)}$, which are independent of $k a$, will be evaluated by finding the roots of Eq. (7.2.44). With those quantities known for each azimuthal harmonic $m$, we may solve Eq.(7.2.45) for all $W_{m} / a$. The right side of that equation is $(a / h) F_{m} /\left(\rho_{\mathrm{e}} c_{\mathrm{e}}^{2}\right)$. In view of Eq. (1), we shall nondimensionalize this term by defining a variable $\chi_{m}$ that replaces $W_{m} / a$, such that

$$
\begin{equation*}
\frac{W_{m}}{a}=\frac{\hat{F}}{\rho_{\mathrm{e}} c_{\mathrm{e}}^{2} a h} \chi_{m} \tag{2}
\end{equation*}
$$

The result of changing variables in this manner is to convert Eq. (7.2.45) to

$$
\begin{align*}
& {\left[\frac{\left[\left(\kappa_{m}^{(1)}\right)^{2}-\left(k_{\mathrm{e}} a\right)^{2}\right]\left[\left(\kappa_{m}^{(2)}\right)^{2}-\left(k_{\mathrm{e}} a\right)^{2}\right]}{\left(\kappa_{m}^{(0)}\right)^{2}-\left(k_{\mathrm{e}} a\right)^{2}}\right.}  \tag{3}\\
& \left.\quad+\frac{\rho_{0} a}{\rho_{\mathrm{e}} h}\left(\frac{c}{c_{\mathrm{e}}}\right)^{2} \frac{k a h_{m}(k a)}{h_{m}^{\prime}(k a)}\right] \chi_{m}=\frac{2 m+1}{4 \pi}
\end{align*}
$$

The values of $\chi_{m}$ may be used to evaluate the coefficients of the pressure series. Introducing Eq. (2) into (7.2.41) leads to

$$
\begin{equation*}
B_{m}=\rho_{0} c^{2} \frac{k a}{h_{m}^{\prime}(k a)} \frac{W_{m}}{a}=\left(\frac{\hat{F}}{a h}\right) \frac{\rho_{0} c^{2}}{\rho_{\mathrm{e}} c_{\mathrm{e}}^{2}} \frac{k a}{h_{m}^{\prime}(k a)} \chi_{m} \tag{4}
\end{equation*}
$$

Thus, the $\chi_{m}$ values found from Eq.(3) are the (dimensional) displacement coefficients ratioed to $\hat{F} /\left(\rho_{\mathrm{e}} c_{\mathrm{e}}^{2} h\right)$. Using those values to evaluate, Eq. (4) yields the pressure coefficients ratioed to $\hat{F} /(a h)$. The displacement and pressure at any location are found by synthesizing the series in Eq. (7.2.35).

Part (a) requested the $W_{m}$ coefficients, which now are $\chi_{m}$, for $m \leq 4$. However, evaluation of the displacement and pressure series requires computations for increasing $m$ until the series converge. Thus, our strategy is to solve Eq. (3) at the full range of frequencies, which we will designate $\left(k_{e} a\right)_{j}$. The result will be a matrix $[\tilde{\chi}]$ whose element at row $m+1$ and column $j$ is $\tilde{\chi}_{m, j}=\chi_{m}$ at $\left(k_{e} a\right)_{j}$. It is logical to follow
the evaluation of the $\chi_{m}$ coefficients with an evaluation of the corresponding $B_{m}$ coefficients. Doing so will yield a matrix $[B]$ that matches $[\chi]$.

Standard properties for aluminum are $\rho_{\mathrm{e}}=2100 \mathrm{~kg} / \mathrm{m}^{3}, c_{\mathrm{e}}=6420 \mathrm{~m} / \mathrm{s}$, and $\nu=0.33$. The in-vacuo natural frequencies are tabulated below.

|  | Spherical harmonic $m$ |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0 | 1 | 2 | 3 | 4 |
| $\Omega_{m}^{(1)}$ | - | 0 | 0.7014 | 0.8327 | 0.8895 |
| $\Omega_{m}^{(2)}$ | 1.6125 | 1.9749 | 2.7221 | 3.6351 | 4.5967 |

Figure 1a shows the frequency dependence of $\left|\chi_{m}\right|$ for $m=0$ to 4 . As expected, each amplitude has a zero at $k_{\mathrm{e}} a=\kappa_{m}^{(0)}$, and it falls off with increasing frequency. Although the theory indicates that each spherical harmonic should show two fluidloaded resonances because there are two natural frequencies at each $m$ other than $m=0$, only one resonance appears in the figure. Figure 1 b provides a zoomed view of this range. It indicates that $\left|W_{0}\right|$ does not feature a strong resonance. The fluidloaded resonance of $W_{1}$ occurs at $k_{e} a=0$. Increasing $m$ beyond one shifts the peak of $\left|W_{m}\right|$ to a higher frequency, with a narrower width and greater height. (In the study of mechanical and electrical oscillations, this feature would be said to be that the $Q$ factor of the resonance increases.) These fluid-loaded resonances occur at substantially lower frequencies than the corresponding in-vacuo natural frequency.


Figure 1a.


Reasonable questions to ask at this juncture are: Why don't any of the curves show two peaks if there are two natural frequencies for each $m>0$ ? Why do peaks occur at
frequencies that are much less than either natural frequency for that $m$ ? Why doesn't the $m=0$ curve have a sharp peak? Why does the $m=1$ curve rise without bound as the frequency approaches zero? The answer to each may be found in the general analysis, but it is helpful to examine them from a slightly different perspective. Rather than examining impedances that relate the applied force to velocity coefficients, let us write Eq. (3) for the displacement coefficients as

$$
\left(K_{\mathrm{e}}+K_{\mathrm{f}}\right) \chi_{m}=\frac{2 m+1}{4 \pi}
$$

The stiffness of an elastic spring is the ratio of the internal force to the elongation. The preceding describes a similar relation, so $K_{\mathrm{e}}$ and $K_{\mathrm{f}}$ are, respectively, called the structural and fluid dynamic stiffnesses. A comparison of the above definition to Eq. (7.2.50) shows that these quantities are related to the respective impedances by

$$
K_{\mathrm{e}}=i k_{e} a Z_{\mathrm{e}}, \quad K_{\mathrm{f}}=i k_{e} a Z_{\mathrm{f}}
$$

We begin with an examination of the structural reactance. The above definitions show that the real part of a dynamic stiffness is $-k_{e} a$ times the reactance. The imaginary part of a dynamic stiffness is $k_{e} a$ times the resistance. Dissipation in the shell material has been neglected, so the sole contributor to $\operatorname{Im}\left(K_{e}+K_{f}\right)$ is the fluid resistance, which means that $\operatorname{Im}\left(K_{\mathrm{e}}+K_{\mathrm{f}}\right)>0$ at any frequency.

Let us examine how the behavior of the dynamic stiffnesses affects each $\chi_{m}$, beginning with $m=2$. Fig. 2 depicts the nonzero parts of $K_{\mathrm{e}}$ and $K_{\mathrm{f}}$. If fluid loading was not present, a zero of $\operatorname{Re}\left(K_{\mathrm{e}}\right)$ would lead to a singular value of $\chi_{m}$. Close inspection shows that this curve does indeed cross the zero axis at $k_{\mathrm{e}} a=\Omega_{2}^{(1)}$ and $\Omega_{2}^{(2)}$. Now suppose that fluid loading was such that $\operatorname{Im}\left(K_{\mathrm{f}}\right)$ were zero. Then, a singular value of $W_{m}$ would occur at any frequency at which $\operatorname{Re}\left(K_{\mathrm{e}}+K_{\mathrm{f}}\right)=0$. The plotted data indicates that this condition occurs at $k_{\mathrm{e}} a=0.276$, which is much less than $\Omega_{2}^{(1)}$, and at $k_{\mathrm{e}} a=2.668$, which is slightly less than $\Omega_{2}^{(2)}$. Of course, $\operatorname{Im}\left(K_{\mathrm{f}}\right)$ is not zero. However, in the vicinity of $k_{\mathrm{e}} a=0.276$, the value of $\operatorname{Im}\left(K_{\mathrm{f}}\right)$ is comparable to the small value of $\operatorname{Re}\left(K_{\mathrm{e}}+K_{\mathrm{f}}\right)$. Consequently, division of the $F_{2}$ force by $K_{\mathrm{e}}+K_{\mathrm{f}}$ yields a large value of $\left|\chi_{2}\right|$. In fact, the maximum is $\left|\chi_{2}\right| / a=10.704$ at $k_{\mathrm{e}} a=0.274$.


Figure 2.

Now consider the situation for $\chi_{2}$ in the vicinity of $k_{\mathrm{e}} a=2.668$, at which $\operatorname{Re}\left(K_{\mathrm{e}}+K_{\mathrm{f}}\right)=0$. The value of $\operatorname{Im}\left(K_{\mathrm{f}}\right)$ in this range is much greater than $\operatorname{Re}\left(K_{\mathrm{e}}+\right.$ $K_{\mathrm{f}}$ ). This means that radiation damping is high, so there is no frequency in this range at which $\left|K_{\mathrm{e}}+K_{\mathrm{f}}\right|$ is noticeably diminished. It is for this reason that $\left|\chi_{2}\right|$ does not feature a peak in this range.

The situation for $m>2$ is similar to what we have seen for $m=2$. The behaviors for $m=0$ and $m=1$ are different because of the special nature of the natural frequencies for each case. Figure 3 displays the dynamic stiffnesses for $m=0$. There is only one natural frequency, so there is only one frequency at which $\operatorname{Re}\left(K_{\mathrm{e}}\right)=0$. The addition of fluid loading leads to $\operatorname{Re}\left(K_{\mathrm{e}}+K_{\mathrm{f}}\right)$ being zero at a slightly lower value of $k_{\mathrm{e}} a$, but $\operatorname{Im}\left(K_{\mathrm{f}}\right)$ is quite large in this range. This means that there is no frequency at which $\left|K_{\mathrm{e}}+K_{\mathrm{f}}\right|$ is greatly reduced, so is there no frequency at which $\left|\chi_{0}\right|$ shows a sharp peak.


Figure 3.

Figure 4 explains the $m=1$ behavior. The lower natural frequency is $\Omega_{1}^{(1)}=0$, and $K_{\mathrm{f}}=0$ at that frequency because low frequencies correspond to low velocities. Hence, $K_{\mathrm{e}}+K_{\mathrm{f}}=0$ at $k_{\mathrm{e}} a=0$, which leads to a singular value of $\left|W_{1}\right| / a$. The behavior in the vicinity of the higher natural frequency is the same as it was for $m=2$. Adding $K_{\mathrm{f}}$ to $K_{\mathrm{e}}$ adds a large imaginary part. Thus, even though there is a frequency at which $\operatorname{Re}\left(K_{\mathrm{e}}+K_{\mathrm{f}}\right)$ is zero, nowhere in that frequency range is $\left|K_{\mathrm{e}}+K_{\mathrm{f}}\right|$ small. The consequence is that large values of $\left|W_{1}\right| / a$ only occur at low frequencies.


Figure 4.

For Part (b), we observe that the displacement at polar angles $\psi_{1}=0$ and $\psi_{2}=$ $\pi$ when the frequency is a specific value $(k a)_{n}$ may be computed by the matrix algorithm. Let [ $L$ ] be a $2 \times(N+1)$ array whose elements are $L_{j, m}=P_{m}\left(\cos \psi_{j}\right)$, that is, row $j$ has the Legendre functions at all orders for $\psi_{j}$. Also, define a rectangular array $[\chi]$ whose columns hold the $\chi_{m}$ values at each $(k a)_{n}$. Then, the displacement series in Eq. (7.2.35) reduces to

$$
\frac{\rho_{\mathrm{e}} c_{\mathrm{e}}^{2} h}{\hat{F}}[w]=[L][\chi]
$$

It is at this juncture that the question of series convergence arises. Given that the $F_{m}$ coefficients in Eq. (2) increase with increasing $m$, it is reasonable to anticipate that the spherical harmonic $M$ at which the summation is halted will be large. If we perform the first calculation with a large $M$, then results for smaller $M$ may be found by deleting columns from the right of $[L]$, and deleting the same number of rows from the bottom of $[\chi]$. Results for $M=100,67$, and 33 are described by Fig. 5. There is no perceptible difference between the results for the two longer series. Thus, it is reasonable to conclude that $M=67$ is adequate for the range $k_{\mathrm{e}} a<8$. The series for $M=33$ agrees well up to $k_{\mathrm{e}} a=4$, then diverges drastically for $k_{\mathrm{e}} a>6$. There is a simple explanation for this behavior. Each $\chi_{m}$, other than $\chi_{0}$, has a narrow peak in the vicinity of a single frequency. Because the displacement $w$ is formed from a weighted series of the $\chi_{m}$ coefficients, $w / a$ may be expected to have peaks at each of these frequencies. The highest peak for $M=33$ occurs at $k_{\mathrm{e}} a=5.908$, which is where the series begins to diverge from the correct solution. The conclusion is that the series length must be sufficiently long to capture all fluid-loaded "resonances" that occur within the frequency range of interest. Unfortunately, this criterion is not much use a priori, because we find these frequencies at part of the analysis.


Figure 5.

A different perspective for convergence results from considering the criterion for accuracy of the shell theory. It was stated as requiring that $m<(\pi / 5) a / h$. For the present system, this requires $m \leq 25$. A graph like Fig. 5 would show that a truncation
of the series at $m=25$ would yield a result that is convergent only for $k_{e} a<3.3$. Given that, we should not assume that our results for higher $k_{e} a$ are what would be observed in an experiment.

The farfield pressure is given by Eq. (7.2.43), which also may be computed as a matrix product. A rectangular array $[C]$ is computed by multiplying each row $\left[B_{m}\right]$ by a phase delay for that $m$, specifically

$$
\left[C_{m}\right]=\left[B_{m}\right] e^{i(m+1) \pi / 2}
$$

The farfield locations at which the pressure is to be evaluated for Part (c) also are $\psi=0$ and $\psi=\pi$, so we may also use [ $L$ ] for this evaluation. Therefore, we evaluate

$$
\left[r P_{\mathrm{ff}, \mathrm{axis}}\right]=[L][C]
$$

Because of the manner in which the $\chi_{m}$ values are defined, the quantities that are evaluated in this manner are $(r / a) P_{\mathrm{ff}}(r, \psi=0$ or $\pi) e^{i k r}(a h / \hat{F})$.

Figure 6 displays the frequency dependence of the farfield pressures on the axis of symmetry. A zoomed view gives a clearer picture of the many low-frequency resonances. The frequencies at which peaks occur are those that maximize one of the $\chi_{m}$ coefficients. An interesting attribute is that the radiation in the forward direction, $\psi=0$, is much greater than the backward direction, $\psi=180^{\circ}$, for nonresonant $k_{\mathrm{e}} a$ values between 0.5 and 5 .


Figure 6.

The directivity requested in Part (d) can be found similarly to the farfield pressure at a specific angle. Let $\left(k_{\mathrm{e}} a\right)_{j}$ denote the frequency of interest, and let $\left\{C_{j}\right\}$ denote the $j$ th column of [C]. The farfield pressure at many angles $\psi_{j}$ may be found by defining a matrix $\left[L^{\prime}\right]$ that differs from $[L]$ only by the fact that its has many rows, $L_{j, m}=P_{m}\left(\cos \psi_{j}\right)$.

A scan of the data shows that the largest value of $r|P|$ for $k_{\mathrm{e}} a<1$ occurs at $k_{e} a=$ 0.404 , and the largest value in the interval $k_{\mathrm{e}} a<5$ occurs at $k_{\mathrm{e}} a=4.935$. Polar plots describing $r|P|$ at these frequencies and $k_{\mathrm{e}} a=2.50$ constitute Fig. 7. The response at the lowest and highest frequencies corresponds to peaks of $\chi_{n}$ for $m=4$ and $m=30$, respectively. Consequently, the farfield pressure at these frequencies is proportional to the Legendre functions $P_{4}(\cos \psi)$ and $P_{30}(\cos \psi)$. In contrast, the middle frequency, $k_{\mathrm{e}} a=2.50$, is not close to any resonance. Hence, all spherical harmonics contribute significantly to the pressure. The consequence is that the directivity plot at $k_{\mathrm{e}} a=2.50$ does not resemble that of a single Legendre function. It will be noted that the overall amplitude at $k_{e} a=2.5$ is comparable to that at the lower frequency, even though $\left|\chi_{4}\right|$ at $k_{e} a=0.404$ is two orders of magnitude greater than any $\left|\chi_{m}\right|$ value at $k_{\mathrm{e}} a=2.5$. This result is a consequence of the fact that the $h_{m}^{\prime}(k a)$ appears in the denominator of Eq. (7.2.43). These values are much larger at the lower frequency. This is another manifestation of the inefficiency of radiation at low frequencies.


Figure 7.

### 7.3 Radiation from an Infinite Cylinder

It is known that the separation of variables approach may be used to solve the Helmholtz equation in eleven orthogonal coordinate systems, such as the spherical coordinates of the previous section. ${ }^{11}$ However, it does not follow that we will always be able to use the technique to analyze radiation from a body whose shape

[^8]Fig. 7.8 An otherwise rigid box undergoing a uniform vibration at velocity amplitude $v_{0}$ along the edge $x=a / 2$

matches one of these coordinate systems. The apparent paradox results from the nature of boundary conditions.

Consider a rectangular box, which is simplified to a two-dimensional system. Suppose the right edge translates, and the other edges are stationary. This is the situation in Fig. 7.8.

The origin of $x y$ has been placed at the center. Thus, the boundary conditions are

$$
\begin{gather*}
\frac{\partial P}{\partial x}=-i \rho_{0} \omega V_{0} \text { on } x=a / 2,-b / 2<y<b / 2 \\
\frac{\partial P}{\partial x}=0 \text { on } x=-a / 2, \quad-b / 2<y<b / 2  \tag{7.3.1}\\
\frac{\partial P}{\partial y}=0 \text { on } y= \pm b / 2, \quad-a / 2<x<a / 2
\end{gather*}
$$

The separation of variables ansatz is that $P=F(x) G(y)$. It is possible to find ordinary differential equations for $F(x)$ and $G(y)$ whose satisfaction assures that $P$ is a solution of the Helmholtz equation. The difficulty is with the boundary conditions. Substitution of the ansatz into the second of the preceding set of boundary conditions gives

$$
\begin{equation*}
\left(\left.\frac{d F}{d x}\right|_{x=-a / 2}\right) G(y)=0,-b / 2<y<b / 2 \tag{7.3.2}
\end{equation*}
$$

In cannot be that $G(y)=0$, so it must be that $d F / d x$ is zero at $x=-a / 2$. However, if it is zero along this edge, then $\partial P / \partial x$ is zero at $x=-a / 2$ for all $y$, not just the edge of the box. In other words, the resulting solution would describe the box plus a rigid sheet at $x=-a / 2$ whose extent is infinite. Similar situations occur for the rigid edges at $y= \pm b / 2$. Thus, the system that is implied by making the separation of variables solution satisfying the rigid boundary conditions is the one depicted in Fig. 7.9. Clearly, this system is not at all like the one we set out to analyze, and it has no physical significance.

Fig. 7.9 System implied by making a separation of variables ansatz satisfying the boundary conditions for the vibrating box in Fig. 7.8. The rigid lines represent rigid thin planes that extend to infinity


Given this observation, it is reasonable to wonder what body shapes can be analyzed for their radiated field by the separation of variables method? If we add the requirement that the analysis not be excessively complicated, and that the system is reasonably common, the answer is: Very few! Spheres were covered in the previous section. Variants featuring ellipses, such as spheroids, have been analyzed, but those studies require great effort because of the nature of the associated separation functions.

Radiation from cylindrical bodies is of great interest because it is an issue that arises in several contexts, including HVAC and automotive exhaust system, and piping and pressure vessel applications. Also, in underwater acoustics, the shape of the pressure hull of a modern submarine usually is a cylinder closed by rounded caps. Unfortunately, the separation of variables solution for a cylindrical geometry does not describe any of these systems. This is so because it is limited to situations where the cylindrical boundary has infinite length. Otherwise, we encounter the same difficulty as that in Fig. 7.9, with a cylindrical boundary extended beyond the ends of the cylinder. Despite this limitation, we will pursue the analysis of infinite cylinders. In part, the justification for doing so is that the investigation will provide results that are good approximations in some regions of the actual field. In addition, analyses like those pursued here often are used to develop simpler models that capture basic phenomena encountered in finite length cylinders.

The system we will explore is described in Fig.7.10. The $z$-axis is defined to coincide with the axis of a cylinder whose radius is $a$ when the cylinder is at rest. The fluid domain extends outward from the cylinder's surface without limit. The surface of the cylinder executes a small amplitude vibration such that its velocity component normal to the surface is $\operatorname{Re}\left(V_{s} e^{i \omega t}\right)$, where $V_{s}$ is an arbitrary specified function of the circumferential angle $\theta$ and axial position $z$. The unit vector $\bar{e}_{R}$ at an arbitrary field point is said to be oriented in the transverse direction relative to the $z$-axis. The circumferential direction is $\bar{e}_{\theta}$, and the axial direction is $\bar{e}_{z}$. Both $\bar{e}_{R}$ and $\bar{e}_{\theta}$ depend on $\theta$, but are independent of $R$ and $z$, whereas $\bar{e}_{z}$ has a fixed orientation.

Fig. 7.10 Coordinate systems and unit vectors for describing a surface vibration of a cylinder


### 7.3.1 Separation of Variables

The pressure field radiated by a vibrating cylinder is a function of $R, \theta$, and $z$. Because $\bar{e}_{R}$ is perpendicular to the surface, the continuity condition requires that the transverse velocity of the fluid equals $V_{s}$. The frequency domain equations governing the field are the Helmholtz equation subject to a boundary condition derived from Euler's equation. The gradient and Laplacian in terms of cylindrical coordinates are derived in Appendix A. The governing equations are

$$
\begin{gather*}
\frac{\partial^{2} P}{\partial R^{2}}+\frac{1}{R} \frac{\partial P}{\partial R}+\frac{1}{R^{2}} \frac{\partial^{2} P}{\partial \theta^{2}}+\frac{\partial^{2} P}{\partial z^{2}}+k^{2} P=0, \quad\left\{\begin{array}{l}
R>a \\
-\pi \leq \theta<\pi \\
-\infty<z<\infty
\end{array}\right.  \tag{7.3.3}\\
\left.\frac{\partial P}{\partial R}\right|_{r=a}=-i \rho_{0} \omega V_{s}, \quad\left\{\begin{array}{l}
-\pi \leq \theta<\pi \\
-\infty<z<\infty
\end{array}\right.
\end{gather*}
$$

We could immediately proceed to apply the method of separation of variables to the Helmholtz equation. However, the circular geometry allows us to simplify identification of the dependence on $\theta$.

If we were to start from any point and follow a circle defined by constant values of $R$ and $z$, an increment of $\theta$ by $2 \pi$ would bring us back to the starting point. It follows that $P$ must be $\theta$-periodic in $2 \pi$, that is, $P(R, \theta, z)=P(R, \theta+2 \pi, z)$, so we introduce a Fourier series. The coefficients of this series may be functions of $R$ and $z$. It is convenient to begin by using a complex series, so we know that

$$
\begin{equation*}
P=\frac{1}{2} \sum_{n=-\infty}^{\infty} F_{n}(R, z) e^{i n \theta} \tag{7.3.4}
\end{equation*}
$$

Because of the linearity of the Helmholtz equation, each term of the Fourier series must constitute a solution. If we substitute a generic term, the harmonic function of $\theta$ appears throughout and may be eliminated. What remains is

$$
\begin{equation*}
\frac{\partial^{2} F_{n}}{\partial R^{2}}+\frac{1}{R} \frac{\partial F_{n}}{\partial R}+\frac{\partial^{2} F_{n}}{\partial z^{2}}+\left(k^{2}-\frac{n^{2}}{R^{2}}\right) F_{n}=0 \tag{7.3.5}
\end{equation*}
$$

At this juncture, we introduce the separation of variable ansatz,

$$
\begin{equation*}
F_{n}=f(R) g(z) \tag{7.3.6}
\end{equation*}
$$

We substitute this form into the differential equation, divide by $f g$, then group terms that are functions of $R$ and $z$ on either side of the equality. The $k^{2}$ term could be placed on either side, but it preferable to keep it with the $1 / R^{2}$ term. Thus, these operations lead to

$$
\begin{equation*}
\frac{1}{f}\left(\frac{d^{2} f}{d R^{2}}+\frac{1}{R} \frac{d f}{d R}\right)+\left(k^{2}-\frac{n^{2}}{R^{2}}\right)=-\frac{1}{g} \frac{d^{2} g}{d z^{2}} \tag{7.3.7}
\end{equation*}
$$

The terms on the left are a function of $R$, while the one on the right depends only on $z$. This can only be true if both sides equal the same constant. The value of this constant is not known a priori, but we do know that $g(z)$ cannot grow as $|z|$ increases. This requires that the separation constant be positive or zero, because a negative value would lead to solution for $g$ that is a growing exponential. Thus, we shall denote the separation constant as $\mu^{2}$. The corresponding separated differential equations are

$$
\begin{gather*}
\frac{d^{2} f}{d R^{2}}+\frac{1}{R} \frac{d f}{d R}+\left(\kappa^{2}-\frac{n^{2}}{R^{2}}\right) f=0 \\
\frac{d^{2} g}{d z^{2}}+\mu^{2} g=0  \tag{7.3.8}\\
\kappa^{2}=k^{2}-\mu^{2}
\end{gather*}
$$

The axial function $g(z)$ is harmonic. In keeping with the usage of a complex Fourier series for the radial distance, this function is written in a similar form as

$$
\begin{equation*}
g=A e^{-i \mu z}+B e^{+i \mu z} \tag{7.3.9}
\end{equation*}
$$

The values of the axial wavenumber $\mu$ and the coefficients $A$ and $B$ are arbitrary in regard to satisfaction of the Helmholtz equation.

### 7.3.2 Transverse Dependence—Cylindrical Bessel Functions

The value of $\mu$ affects $f(R)$ by altering the constant $\kappa$. To begin, we will consider $\mu$ to be less than $k$, so that $\kappa$ is a real number, which may be taken to be positive. The case where $\mu>k$ is important and will be studied after the primary development.

The differential equation for $f(R)$ is reminiscent of the equation governing the radial function for spherical waves. However, the absence of a two factor in the second term, $(1 / R) d f / d R$, is important. It is possible to derive a transformation that converts the present equation to the spherical form. An alternative would derive the solution by following the method of Frobenius, which modifies a power series in $R$ with a term that is an unspecified fractional power. Some readers might have encountered a differential equation like this in other subjects. In any event, the present endeavor is well served by merely observing that the first of Eq.(7.3.8) is Bessel's equation. Its solutions are Bessel and Neumann functions. We encountered them in the previous section, where it was noted there that standard algorithms compute these functions as intermediate steps in the evaluation of spherical Bessel functions.

A rigorous terminology would refer to these functions as cylindrical Bessel functions, but their occurrence is so common that the adjective "cylindrical" usually is dropped. To distinguish them from the spherical functions, capital letters are used for
the cylindrical functions: $J$ for the Bessel function, $N$ for the Neumann (although $Y$ also is often used), and $H$ for the Hankel functions that we soon will define. The value of $n$, which is the order of the function, appears in the subscript. Furthermore, if we were to divide the first of Eq. (7.3.8) by $\kappa$, we would see that the independent variable for this form of Bessel's equation is $\kappa R$, so that is the argument of the functions. Among the many ways to evaluate Bessel functions, the least useful is from their definitions. Nevertheless, they are listed here for the sake of completeness,

$$
\begin{gather*}
J_{n}(x)=\sum_{k=0}^{\infty} \frac{(-1)^{k}}{k!(n+k)!}\left(\frac{z}{2}\right)^{2 k+n} \\
N_{n}(x)=\frac{2}{\pi} \ln \left(\frac{z}{2}\right) J_{n}(x)-\frac{1}{\pi} \sum_{k=0}^{n-1} \frac{(n-k-1)!}{k!}\left(\frac{z}{2}\right)^{2 k-n}  \tag{7.3.10}\\
-\frac{1}{\pi} \sum_{k=0}^{\infty} \frac{(-1)^{k}}{k!(n+k)!}[\Psi(k+1)+\Psi(n+k+1)]\left(\frac{z}{2}\right)^{2 k+n}
\end{gather*}
$$

In the second expression, $\Psi()$ is the psi or digamma function, which is defined in standard references on special functions. ${ }^{12}$

Both $J_{n}(x)$ and $N_{n}(x)$ are real. Therefore, neither by itself can represent a wave propagating in the transverse direction. For spherical waves, the spherical Hankel functions fit this specification. The same is true here. Hankel functions of the first and second kind and order $n$ are defined to be

$$
\begin{equation*}
H_{n}^{(2)}(x)=H_{n}^{(1)}(x)^{*}=J_{n}(x)-i N_{n}(x) \tag{7.3.11}
\end{equation*}
$$

As was true for spherical waves, the second Hankel function represents an outgoing wave at large $k R$ corresponding to our representation of harmonic response as $\operatorname{Re}\left(P e^{i \omega t}\right)$. This fact is seen when we consider the asymptotic behavior of each function as its argument $\kappa R$ increases,

$$
\left.\begin{array}{c}
H_{n}^{(2)}(\kappa R)=H_{n}^{(1)}(\kappa R)^{*} \rightarrow\left(\frac{2}{\pi \kappa R}\right)^{1 / 2} e^{-i(\kappa R-(2 n+1) \pi / 4)}  \tag{7.3.12}\\
J_{n}(\kappa R) \rightarrow\left(\frac{2}{\pi \kappa R}\right)^{1 / 2} \cos \left(\kappa R-\frac{2 n+1}{4} \pi\right) \\
N_{n}(\kappa R) \rightarrow\left(\frac{2}{\pi \kappa R}\right)^{1 / 2} \sin \left(\kappa R-\frac{2 n+1}{4} \pi\right)
\end{array}\right\} \text { as } \kappa R \rightarrow \infty
$$

The Bessel and Neumann functions will arise when we explore cylindrical cavities and waveguides, but we will have no need for the first kind Hankel function. Thus,

[^9]we will drop the superscript "two" notation for acoustical analyses, as we did for spherical waves, so that $H_{n}(k R)$ will always refer to the second Hankel functions.

The behavior of the cylindrical Bessel functions at small arguments also is important. The Bessel function is finite at $x=0$, but the Neumann and Hankel functions are singular. The trends are

$$
\left.\begin{array}{c}
H_{n}^{(2)}(x)=H_{n}^{(1)}(x)^{*} \rightarrow\left\{\begin{array}{l}
-\frac{2 i}{\pi} \ln (x) \text { if } n=0 \\
\frac{i}{\pi}(n-1)!\left(\frac{2}{x}\right)^{n} \text { if } n \geq 1
\end{array}\right\} \\
J_{n}(x) \rightarrow \frac{1}{n!}\left(\frac{x}{2}\right)^{n} \tag{7.3.13}
\end{array}\right\}
$$

The limiting behaviors for small and large arguments are evident in Fig.7.11. The Bessel functions are finite at $x=0$, with all except the zero order being zero there. In contrast, all Neumann functions are negatively infinite at the origin, with a growth rate that increases with increasing order. For large arguments, a Neumann function resembles the Bessel function at that order with a lag of $\pi / 2$. Both functions decay with increasing values of their argument at fixed order. However, the decrease is inversely proportional to the square root of their argument, which is a slower decay than the decay rate for the spherical Bessel functions.


Fig. 7.11 Low-order Bessel and Neumann functions as functions of $\kappa R$

The order $n$ affects the magnitude of the argument $x$ at which it is appropriate to employ Eq. (7.3.12). Specifically, the large argument trends will be observed if $x>n+2 \pi$. The opposite trend is the behavior when the order is much greater than the argument. In this case, we have

$$
\left.\begin{array}{rl}
J_{n}(\kappa R) & \rightarrow\left(\frac{1}{2 \pi n}\right)^{1 / 2}\left(\frac{e \kappa R}{2 n}\right)^{n}  \tag{7.3.14}\\
N_{n}(\kappa R) \rightarrow & -\left(\frac{2}{\pi n}\right)^{1 / 2}\left(\frac{e \kappa R}{2 n}\right)^{-n}
\end{array}\right\} \text { if } n \gg \kappa R
$$

Most standard mathematics software contains routines for evaluating the Bessel and Neumann functions. We also will need to evaluate their derivatives. As always, a prime denotes differentiation of a function with respect to its argument. The following recurrence relations resemble those for spherical Bessel functions. Those for the derivatives will be useful when it is necessary to satisfy velocity boundary conditions, whereas the relations featuring functions at different orders are a core component of most numerical algorithms by which the functions are evaluated. In the following, $F_{n}(x)$ represents any of the cylindrical functions.

$$
\begin{gather*}
\frac{2 n}{x} F_{n}(x)=F_{n-1}(x)+F_{n+1}(x) \\
F_{n}^{\prime}(x)=\frac{1}{2}\left[F_{n-1}(x)-F_{n+1}(x)\right] \\
=F_{n-1}(x)-\frac{n}{x} F_{n}(x) \\
=-F_{n+1}(x)+\frac{n}{x} F_{n}(x)  \tag{7.3.15}\\
F_{0}^{\prime}(x)=-F_{1}(x) \\
\frac{F_{-n}(x)=(-1)^{n} F_{n}(x)}{d x^{j}} F_{n}(x)=\frac{1}{2^{j}}\left[F_{n-j}(x)-\frac{j!}{1!(j-1)!} F_{n-j+1}(x)\right. \\
\left.+\frac{j!}{2!(j-2)!} F_{n-j+2}(x)-\cdots+(-1)^{j} F_{n+j}(x)\right]
\end{gather*}
$$

Many other relations featuring the various Bessel functions are available. ${ }^{13}$

### 7.3.3 Radiation Due to a Helical Surface Wave

The first case we shall consider is that in which the surface motion consists of a single term in the complex Fourier series for the circumferential dependence, and

[^10]Fig. 7.12 Depiction of a helical wave vector $\bar{k}_{s}$. a Unwrapped cylindrical surface. b Tangent to the helix


the $z$ dependence is a sinusoidal wave that propagates in the positive direction. In other words, the radial velocity on the surface is taken to be

$$
\begin{equation*}
v_{s}=\operatorname{Re}\left(V_{0} e^{i(\omega t-\mu z-n \theta)}\right) \tag{7.3.16}
\end{equation*}
$$

Clearly, this represents a wave, but in what direction is it propagating? To answer this, note that the arclength along the cylinder in the circumferential direction is $s=a \theta$. Let us unwrap the surface of the cylinder and lay it on a plane, as is shown in Fig. 7.12a. Position along this flat surface is $\bar{x}=z \bar{e}_{z}+s \bar{e}_{\theta}$, where the unit vectors are those in Fig. 7.10. This suggests that we rewrite the phase variable of $v_{s}$ as $\omega t-\bar{k}_{s} \cdot \bar{x}$, where the surface wavenumber vector is

$$
\begin{equation*}
\bar{k}_{s}=\mu \bar{e}_{z}+\frac{n}{a} \bar{e}_{\theta} \tag{7.3.17}
\end{equation*}
$$

This vector forms a constant angle with $\bar{e}_{z}$, which means that on this plane the wave follows a straight line. When the plane is wrapped back onto the cylinder in Fig. 7.12b, this straight line becomes a helix. Correspondingly, Eq. (7.3.16) is said to be a helical surface wave. In the case where $n=0$, the surface wave is axisymmetric and the wavenumber vector is parallel to the axis of the cylinder.

In addition to the boundary condition at the cylinder's surface, the acoustic response must be an outgoing wave having suitably decaying amplitude as $R$ increases. The only radial dependence fitting this specification with our use of the $e^{i \omega t}$ convention is the Hankel function of the second kind. Furthermore, because $\bar{V} \cdot \bar{e}_{R}$ is obtained from a derivative of $P$ with respect to $R$, the manner in which pressure depends on $z$ and $\theta$ must match the dependence of $V_{s}$. Thus, we shall try

$$
\begin{equation*}
P=B H_{n}(\kappa R) e^{-i(\mu z+n \theta)}, \quad \kappa=\left(k^{2}-\mu^{2}\right)^{1 / 2} \tag{7.3.18}
\end{equation*}
$$

The boundary condition is

$$
\begin{equation*}
\left.\left.\nabla P \cdot \bar{e}_{R}\right|_{R=a} \equiv \frac{\partial P}{\partial R}\right|_{R=a}=-i \omega \rho_{0} V_{s}=-i \omega \rho_{0} V_{0} e^{-i(\mu z+n \theta)} \tag{7.3.19}
\end{equation*}
$$

Our ansatz will satisfy this condition if

$$
\begin{equation*}
B \kappa H_{n}^{\prime}(\kappa a)=-i \omega \rho_{0} V_{0} \equiv-i k\left(\rho_{0} c V_{0}\right) \tag{7.3.20}
\end{equation*}
$$

The corresponding pressure and particle velocity expressions are

$$
\begin{gather*}
P=-i \rho_{0} c V_{0} \frac{k H_{n}(\kappa R)}{\kappa H_{n}^{\prime}(\kappa a)} e^{-i(\mu z+n \theta)} \\
\bar{V}=-\frac{1}{i \omega \rho_{0}}\left(\frac{\partial P}{\partial R} \bar{e}_{R}+\frac{1}{R} \frac{\partial P}{\partial \theta} \bar{e}_{\theta}+\frac{\partial P}{\partial z} \bar{e}_{z}\right)  \tag{7.3.21}\\
=V_{0}\left[\frac{H_{n}^{\prime}(\kappa R)}{H_{n}^{\prime}(\kappa a)} \bar{e}_{R}-i\left(\frac{n}{R} \bar{e}_{\theta}+\mu \bar{e}_{z}\right) \frac{H_{n}(\kappa R)}{\kappa H_{n}^{\prime}(\kappa a)}\right] e^{-i(\mu z+n \theta)}
\end{gather*}
$$

An important aspect of this expression is that $n$ may be positive, negative, or zero. Because of the way the Hankel, Bessel, and Neumann functions are defined, functions of negative order are related to the positive order functions by an alternating sign, specifically

$$
\begin{equation*}
H_{-n}(x)=(-1)^{n} H_{n}(x), \quad J_{-n}(x)=(-1)^{n} J_{n}(x), \quad N_{-n}(x)=(-1)^{n} N_{n}(x) \tag{7.3.22}
\end{equation*}
$$

The sign change has no effect on Eq. (7.3.21) because the terms are ratios of a Hankel function and its derivative at the same order.

The farfield decay of a cylindrical wave is different from the spherical spreading of the field that radiates from a finite body. To see why, suppose we select a field point at a very large value of $R$. The line from the field point to points in the cylinder cross section at the same $z$ is perpendicular to the cylinder's axis. In contrast, lines from the field point to the distant ends of the cylinder are essentially parallel to the $z$-axis. There is no radial distance at which all lines from the field point to the surface are parallel, so the prior theorems regarding farfield behavior do not apply.

To identify the farfield limit, we employ the first of Eq. (7.3.12) to write

$$
\begin{equation*}
P_{\mathrm{ff}}=-i \rho_{0} V_{0} \frac{k e^{i(2 n+1) \pi / 4}}{\kappa H_{n}^{\prime}(\kappa a)}\left(\frac{2}{\pi \kappa R}\right)^{1 / 2} e^{-i(\kappa R+n \theta+\mu z)} \tag{7.3.23}
\end{equation*}
$$

The particle velocity in the farfield is obtained by substituting $P_{\mathrm{ff}}$ into Euler's equation. The fact that $R$ is large renders the circumferential velocity negligible compared to $\bar{V} \cdot \bar{e}_{R}$ and $\bar{V} \cdot \bar{e}_{z}$. The result is

$$
\begin{equation*}
\bar{V}_{\mathrm{ff}}=\frac{P_{\mathrm{ff}}}{\rho_{0} c}\left(\frac{\kappa}{k} \bar{e}_{R}+\frac{\mu}{k} \bar{e}_{z}\right) \tag{7.3.24}
\end{equation*}
$$

The relation of pressure and particle velocity in a plane wave is $\bar{V}=\left(P / \rho_{0} c\right) \bar{k} / k$. The above expression suggests a similar form. The phase variable of the farfield pressure, Eq. (7.3.23), is $\kappa R+n \theta+\mu z$. It may be expressed in terms of a wavenumber vector $\bar{k}^{\prime}$ if we let $\bar{x}$ be the position off the cylinder after it is unwrapped, that is, $\bar{x}=R \bar{e}_{r}+s \bar{e}_{\theta}+z \bar{e}_{z}$. Then, setting $\bar{k}^{\prime} \cdot \bar{x}=\kappa R+n \theta+\mu z$ leads to

$$
\begin{equation*}
\bar{k}^{\prime}=\kappa \bar{e}_{r}+\frac{n}{R} \bar{e}_{\theta}+\mu \bar{e}_{z} \tag{7.3.25}
\end{equation*}
$$

Fig. 7.13 Wavenumber vectors for the acoustic wave generated by a helical wave on the surface of a cylinder


Figure 7.13 shows $\bar{k}^{\prime}$ at two locations along a transverse line, as well as $\bar{k}_{s}$ for the surface wave. (In a strict sense, we should not depict $\bar{k}^{\prime}$ close to the surface unless $\kappa a \gg 1$, because Eq. (7.3.23) is only valid if $\kappa R \gg 1$.)

At the surface, the portion of $\bar{k}^{\prime}$ that is parallel to the surface, that is the $\bar{e}_{\theta}$ and $\bar{e}_{z}$ components, matches the surface wavenumber $\bar{k}_{s}$. In other words, the trace wavenumbers match, and the surface wave is supersonic. At very large $R$, the $\bar{e}_{\theta}$ component of $\bar{k}^{\prime}$ is negligible. Thus, Eq. (7.3.24) states that

$$
\begin{equation*}
\bar{V}_{\mathrm{ff}}=\frac{P_{\mathrm{ff}}}{\rho_{0} c} \lim _{R \rightarrow \infty}\left(\frac{\bar{k}^{\prime}}{k}\right) \tag{7.3.26}
\end{equation*}
$$

Furthermore, it follows from the definition of $\kappa$ that $\left|\bar{k}^{\prime}\right| \rightarrow k$ as $R \rightarrow \infty$. In other words, in the farfield, the wave propagates in the plane formed by the field point and the cylinder's axis. Over short propagation distances, it seems to be locally planar, but cylindrical spreading is observable for large changes in $R$.

Whether a surface wave results in radiation depends only on the scale of the axial dependence. The condition that $\mu<k$ for a wave to radiate to the farfield is equivalent to saying that the axial wavelength $2 \pi / \mu$ is greater than the acoustic wavelength $2 \pi / k$, or equivalently, that the axial propagation speed $c_{s}$ exceeds the speed of sound, $c \equiv \omega / k$. Each is a manifestation of trace matching. It is evident that the circumferential component of $\bar{k}_{s}$ is not relevant to the matching of trace velocities. In essence, as waves propagate along the cylinder, the axial features do not change, whereas the surface falls away in the circumferential direction. Thus, the circumferential effect is local, whereas the axial effect is global. (This is the reason why the phenomenon of matching trace velocity is not evident in radiation from spheres.)

## Subsonic Surface Waves-Modified Bessel Functions

Recognition that the axial trace velocity of the acoustic wave must equal the axial speed of the surface wave leads to the same question as that for a surface wave on a
plane, specifically, what if the surface wave is subsonic? The axial component of the surface wave's phase velocity is $\omega / \mu$. The supersonic case is obtained if $\mu<k$, in which case, the value of $\kappa^{2}$ is positive. Conversely, a subsonic wave corresponds to $\mu>k$, in which case, $\kappa^{2}$ is negative. When this situation occurs on a planar boundary, the result is an evanescent acoustic wave. The same is true here, but the fact that the dependence on $R$ is described by a Hankel function is a slight complication.

To make it clear that $\kappa^{2}$ is negative, we replace it with $-\beta^{2}$, with $\beta>0$. Then, the differential equation resulting from the separation of variables procedure is

$$
\begin{equation*}
\frac{d^{2} f}{d R^{2}}+\frac{1}{R}-\left(\beta^{2}+\frac{n^{2}}{R^{2}}\right) f=0, \quad \beta^{2}=\mu^{2}-k^{2} \tag{7.3.27}
\end{equation*}
$$

This is a modified Bessel's equation, and its solutions are the real functions $K_{n}(\beta R)$ and $I_{n}(\beta R)$, which, respectively, are the modified Bessel functions of the first and second kind. The $K_{n}(\beta R)$ function is singular at the origin and decays to zero as $\beta R \rightarrow \infty$. In contrast, $I_{n}(\beta R)$ is finite at $\beta R=0$ and grows without bound as $\beta R \rightarrow \infty$. The dependence of the lowest order functions is depicted in Fig. 7.14.


Fig. 7.14 Modified Bessel functions for orders zero to four
Asymptotic approximations of the modified functions are

$$
\left.\left.\begin{array}{rl}
K_{0}(x) & \rightarrow-\ln (x), \quad K_{n}(x) \tag{7.3.28}
\end{array}\right) \frac{1}{2}(n+1)!\left(\frac{2}{x}\right)^{n}\right) \quad \text { as } x \rightarrow 0
$$

The recurrence relations for these functions are slightly different from those for the regular functions. In the following, $F_{m}(x)$ is either $I_{n}(x)$ or $(-1)^{n} K_{n}(x)$,

$$
\begin{gather*}
\frac{2 n}{x} F_{n}(x)=F_{n-1}(x)-F_{n+1}(x) \\
\begin{aligned}
F_{n}^{\prime}(x) & =\frac{1}{2}\left[F_{n-1}(x)+F_{n+1}(x)\right] \\
& =F_{n-1}(x)-\frac{n}{x} F_{n}(x) \\
& =F_{n+1}(x)+\frac{n}{x} F_{n}(x) \\
K_{0}^{\prime}(x) & =-K_{1}(x), \quad I_{0}^{\prime}(x)=I_{1}(x)
\end{aligned}
\end{gather*}
$$

Although both $I_{n}(\beta R)$ and $K_{n}(\beta R)$ satisfy Eq. (7.3.27), the former grows exponentially with increasing $R$. This behavior is counter to the Sommerfeld radiation condition, so it is discarded. The modified function of the second kind decays exponentially as $R$ increases, so it represents an evanescent wave. Thus, the pressure generated from a high wavenumber on the surface is

$$
\begin{equation*}
P_{\mathrm{ev}}=B K_{n}(\beta R) e^{-i(\mu z+n \theta)}, \quad \mu>k \tag{7.3.30}
\end{equation*}
$$

The corresponding particle velocity is

$$
\begin{equation*}
\bar{V}=-\frac{B}{i \omega \rho_{0}}\left[\beta K_{n}^{\prime}(\beta R) \bar{e}_{R}-i K_{n}(\beta R)\left(\frac{n}{R} \bar{e}_{\theta}+\mu \bar{e}_{z}\right)\right] e^{-i(\mu z+n \theta)} \tag{7.3.31}
\end{equation*}
$$

The transverse component must match the surface velocity in Eq. (7.3.16), which sets $B$. When that expression is substituted into the above general solution, the result is

$$
\begin{equation*}
P_{\mathrm{ev}}=-i \rho_{0} c V_{0} \frac{k K_{n}(\beta R)}{\beta K_{n}^{\prime}(\beta a)} e^{-i(\mu z+n \theta)} \tag{7.3.32}
\end{equation*}
$$

The fact that subsonic surface waves do not radiate to the farfield is the same property as subsonic waves on the planar boundary of a half-space. Another aspect becomes evident when we recognize that the modified Bessel functions are real. This means that the transverse component of velocity, $\bar{V} \cdot \bar{e}_{R}$, is everywhere $90^{\circ}$ out-of phase from the pressure. One consequence is that power in not radiated when $\mu>k$. It also has implications for fluid-structure interaction, which is touched on in the next example.

In some applications, such as computer programs, it is inconvenient to monitor whether $\mu>k$, and then select the appropriate description of $P$. In the case of waves on a planar surface, we found that we could use the equations for supersonic waves to describe the subsonic case. With the objective of doing so here, we note that the
modified functions are related to the regular Bessel functions, which should not be surprising in view of the fact that the equations they satisfy only differ by a sign. The relations are

$$
\begin{gather*}
I_{n}(x)=(-i)^{n} J_{n}(i x) \\
K_{n}(x)=\frac{\pi}{2}(-i)^{n+1} H_{n}(-i x) \tag{7.3.33}
\end{gather*}
$$

In the present context, the values of $x$ of interest are $\beta a$ and $\beta R$. By definition, $\kappa= \pm i \beta$. If we use the positive sign, then $H_{n}(\kappa R)$ will become $H_{n}(i \beta R)$, in which the above relation indicates to be proportional $K_{n}(-\beta R)$. We wish that $H_{n}(\kappa R)$ convert to $K_{n}(+\beta R)$, so we set $\kappa=-i \beta$ for subsonic surface waves. Thus, we may employ the supersonic solution, Eq. (7.3.21), to treat any helical surface wave, provided that the transverse wavenumber is set according to

$$
\kappa=\left\{\begin{array}{l}
\left(k^{2}-\mu^{2}\right)^{1 / 2} \text { if } \mu<k  \tag{7.3.34}\\
-i\left(\mu^{2}-k^{2}\right)^{1 / 2} \text { if } \mu>k
\end{array}\right.
$$

Of course, this scheme is only useful if the computational routine that evaluates Hankel functions allows the argument to be imaginary.

The alteration from a regular to a modified Bessel function has an important influence on the radiation impedance, which is a measure of how effectively a surface motion generates a pressure wave. The impedance of a locally reacting surface is the ratio of the surface pressure to the inward normal surface velocity. In contrast, the radiation impedance is the proportionality of the surface acoustic pressure resulting from a known outward surface velocity. The relation is $Z_{\text {rad }}=P / \bar{V}_{s} \cdot \bar{n}$. Despite the apparent similarity in their definitions, the two types of impedances are fundamentally different. A local impedance is a property of the surface, such as its viscoelastic properties. The radiation impedance is an acoustical property that depends on the shape of the radiating body and the nature of the surface motion. If we know it as a function of the fundamental parameters of the system, such as the wavenumbers and frequencies of the surface motion, it may be used to describe the behavior in a variety of situations.

The radiation impedance of a helical wave is obtained by evaluating the pressure at $R=a$, then dividing it by the surface velocity. The velocity is given by Eq. (7.3.16). The surface pressure in the case of a supersonic propagation in the axial direction is found by setting $R=a$ in Eq.(7.3.21), whereas setting $R=a$ in Eq.(7.3.32) gives the surface pressure in the subsonic case. The radiation impedance that results is given by

$$
\frac{Z_{\mathrm{rad}}}{\rho_{0} c}=\left\{\begin{array}{l}
-i \frac{k a}{\kappa a} \frac{H_{n}(\kappa a)}{H_{n}^{\prime}(\kappa a)} \text { if } \mu a<k a  \tag{7.3.35}\\
-i \frac{k a}{\beta a} \frac{K_{n}^{n}(\beta a)}{K_{n}^{\prime}(\beta a)} \text { if } \mu a>k a
\end{array}\right.
$$

The nondimensional form of this expression tends to facilitate numerical evaluations, which are conducted in the next example.

A fundamental difference between the supersonic and subsonic cases is evident in this expression for $Z_{\text {rad }}$. The regular Hankel function is complex, which means that $Z_{\text {rad }}$ has both a resistive and a reactive component. The resistive part is associated with transfer of power from the surface into the fluid. In contrast, the modified Bessel functions are real, so $Z_{\text {rad }}$ is imaginary for subsonic waves. In other words, it is purely a reactance. Because the pressure is $90^{\circ}$ out-of-phase from the velocity, the time-averaged power flow into the fluid will be zero.

EXAMPLE 7.6 The task here is to determine the radiation impedance associated with a helical surface wave. Evaluate and graph the result as a function of $\mu a$, with $k a=2$ and 10 , and $n=0,2$, and 4 .

## Significance

If the radiation impedance is known at the outset of the analysis, it can be used to describe the pressure effect on the structure without performing a parallel acoustical analysis. The parameter dependencies that will be disclosed in the course of this example have much significance.

## Solution

How we proceed depends on whether our computational software can evaluate Hankel function for an imaginary argument. If so, then the field for any set of wavenumbers may be computed by employing Eq. (7.3.34) in conjunction with the first of Eq. (7.3.35) to evaluate $Z_{\text {rad }}$ for any value of $\mu$. If not, then both parts of Eq. (7.3.35) must be constructed explicitly. A vector $\{\mu a\}$ that covers the full range of axial wavenumbers $m$ consists of two partitions. The upper vector $\{\kappa a\}$ corresponds to values of $m$ for which $\mu_{m} a>\kappa a$, so its elements are defined according to Eq. (7.3.18). The lower vector $\{\beta a\}$ corresponds to $\mu_{m} a>\kappa a$, so Eq. (7.3.27) defines its elements. A vector $\left\{Z_{\mathrm{rad}}\right\}$ then is obtained by stacking the values obtained from $\{\beta a\}$ below those obtained from $\{\kappa a\}$.

The Hankel function is complex, so $Z_{\text {rad }}$ for supersonic waves is complex. In contrast, the modified Bessel functions are real and positive, and their derivative is real and negative. Therefore, $Z_{\text {rad }}$ for subsonic waves is always positive imaginary, that is, it is an inertance. Figure 1 shows the low-frequency behavior. There is a singularity at $\mu=k$ for $n=0$, but not for $n>0$. This difference stems from the behavior of the Bessel functions as their argument approaches zero. For $n=0, H_{0}(\kappa a)$ has a logarithmic singularity as $\kappa a \rightarrow 0$. Equation (1) indicates that $Z_{\mathrm{rad}}$ in that case is proportional to $\ln (\kappa a)$. The behavior for $\beta a \rightarrow 0$ is similar. In contrast, for $n>0$, the trend is that $H_{n}(\kappa a)$ is proportional to $(\kappa a)^{-n}$ as $\kappa a \rightarrow 0$. The result is that $Z_{\text {rad }}$ is finite in the limit. Another interesting aspect is that $\operatorname{Im}\left(Z_{\mathrm{rad}}\right)>0$ for all values of $\mu$, which means that its reactive part always is an inertance. We also see that the impedance at any value of $\mu a$ decreases as $n$ increases.


Figure 1.

Increasing the frequency changes some of these trends. The singular behavior in Fig. 2 around $\mu=k$ is like Fig. 1, but both parts of $Z_{\mathrm{rad}}$ for $n>0$ are maximized slightly below $\mu=k$. Also, the intertances above $\mu=k$ are greater than they were at the lower frequency.


Figure 2.

From the viewpoint of the structure, a large value of $\left|Z_{\text {rad }}\right|$ means that any movement will be strongly opposed by the fluid. Larger values of $\left|Z_{\mathrm{rad}}\right|$ in the vicinity of $\mu=k$ mean that it would be more difficult to coerce the structure to move in the transition from supersonic to subsonic surface waves.

### 7.3.4 Axially Periodic Surface Vibration

A helical wave is a building block from which more complicated patterns of surface vibration may be synthesized. How that synthesis is implemented depends on the extent of the vibration along the axis of the cylinder. The simpler situation is that of an axially periodic vibration. In this case, the surface velocity may be described by combining a Fourier series for the $z$ dependence with a Fourier series in $\theta$. This
pattern is replicated by the pressure. The more difficult situation to analyze is that in which the surface vibration covers a finite length of the cylinder, with the remainder of the surface being quiescent. That case is taken up in the next section.

Equation (7.3.4) described the surface velocity as a Fourier series composed of circumferential harmonics. Our interest here is the situation where the surface velocity also is periodic over an axial distance $L$. Such a condition requires that the coefficient of each circumferential harmonic has that period in the axial direction. Thus, the surface velocity may be described by a double Fourier series, with coefficients denoted as $V_{m, n}$ to indicate the associated axial harmonic $m$ and circumferential harmonic $n$. Hence, the surface velocity in the transverse direction is represented as

$$
\begin{equation*}
V_{s}=\frac{1}{4} \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} V_{m, n} e^{-i(2 m \pi x / L+n \theta)} \tag{7.3.36}
\end{equation*}
$$

This velocity distribution is a superposition of helical surface waves whose wavenumbers are $\mu=2 m \pi / L$ in the $z$ direction and $n$ in the $\theta$ direction. The corresponding pressure is obtained by adding the contribution associated with each helical wave, which is given by the first of Eq. (7.3.21) for supersonic waves, and Eq. (7.3.32) for subsonic surface waves. To distinguish the two in the summation, we define a cutoff axial harmonic that marks the highest axial harmonic for which $\mu<k$. This value is

$$
\begin{equation*}
M=\text { floor }\left(\frac{k L}{2 \pi}\right) \tag{7.3.37}
\end{equation*}
$$

The axial harmonics for $m \leq M$ are supersonic, so they form the radiated pressure $P_{\mathrm{rad}}$, while the higher harmonics form the evanescent field $P_{\mathrm{ev}}$,

$$
\begin{align*}
& P=P_{\mathrm{rad}}+P_{\mathrm{ev}} \\
& P_{\mathrm{rad}}=-\frac{1}{4} i \rho_{0} c \sum_{n=-\infty}^{\infty} \sum_{m=-M}^{M} V_{m, n} \frac{k H_{n}\left(\kappa_{m} R\right)}{\kappa_{m} H_{n}^{\prime}\left(\kappa_{m} a\right)} e^{-i 2 m \pi z / L} e^{-i n \theta} \\
& P_{\mathrm{ev}}=-\frac{1}{4} i \rho_{0} c \sum_{n=-\infty}^{\infty} \sum_{m=M+1}^{\infty}\left[V_{m, n} e^{-i 2 m \pi z / L}+V_{(-m), n} e^{i 2 m \pi z / L}\right] \frac{k K_{n}\left(\beta_{m} R\right)}{\beta K_{n}^{\prime}\left(\beta_{m} a\right)} e^{-i n \theta} \tag{7.3.38}
\end{align*}
$$

A subscript $m$ has been assigned to the wavenumbers of the Bessel functions to indicate that they depend on the axial harmonic number. Specifically,

$$
\begin{equation*}
\kappa_{m}=\left(k^{2}-\frac{4 m^{2} \pi^{2}}{L^{2}}\right)^{1 / 2}, \quad \beta_{m}=\left(\frac{4 m^{2} \pi^{2}}{L^{2}}-k^{2}\right)^{1 / 2} \tag{7.3.39}
\end{equation*}
$$

The signal generated by a subsonic surface wave does not reach the farfield. The farfield representation is obtained by dropping $P_{\mathrm{ev}}$ and applying Eq.(7.3.12) to approximate $H_{n}\left(\kappa_{n} R\right)$. This leads to

$$
\begin{align*}
P_{\mathrm{ff}} & =-\frac{1}{4} i \rho_{0} c \sum_{n=-\infty}^{\infty} \sum_{m=-M}^{M} \frac{k V_{m, n}}{\kappa_{m} H_{n}^{\prime}\left(\kappa_{m} a\right)}\left(\frac{2}{\pi \kappa_{m} a}\right)^{1 / 2}\left(\frac{a}{R}\right)^{1 / 2} e^{i(2 n+1) \pi / 4}  \tag{7.3.40}\\
& \times e^{-i\left(\kappa_{m} R+n \theta+2 m \pi z / L\right)}
\end{align*}
$$

Equation (7.3.36) describes $V_{s}$ as a superposition of helical waves that propagate in the positive $z$ direction $(m>0)$ and negative $z$ direction $(m<0)$, with each set consisting of waves that twist in the direction of increasing $\theta(n>0)$ and decreasing $\theta(n<0)$. Some individuals prefer to describe the dependencies on $\theta$ and $z$ as a superposition of standing waves, rather than propagating helical waves. Such a form may be obtained by applying Euler's identity to the complex exponentials. The result would be four terms that are either $\cos (n \theta)$ or $\sin (n \theta)$ multiplied by either $\cos (2 m \pi x / L)$ or $\sin (2 m \pi x / L)$.

It is sufficient to determine the time-averaged power radiated from a segment of the cylinder whose length is $L$ because the spatial field is replicated over this distance. An evaluation of radiated power begins by surrounding the radiating cylinder with an infinitely long concentric cylindrical surface. The radius of the surrounding cylinder is taken to be sufficiently large that the farfield approximation applies. The outward normal to the surrounding cylinder is $\bar{e}_{R}$, so we must construct $\left(I_{R}\right)_{\mathrm{av}}$. The farfield pressure is described by Eq. (7.3.40). When the terms in that series are denoted as $\left(P_{\mathrm{ff}}\right)_{m, n}$, the particle velocity for each term is described by Eq. (7.3.24). The transverse velocity component is associated with radiation, so we have

$$
\begin{equation*}
\left(I_{R}\right)_{\mathrm{av}}=\frac{1}{2 \rho_{0} c} \sum_{n=-\infty}^{\infty} \sum_{m=-M}^{M} \sum_{j=-\infty}^{\infty} \sum_{\ell=-M}^{M} \operatorname{Re}\left[\frac{\kappa_{m}}{k}\left(P_{\mathrm{ff}}\right)_{m, n}\left(P_{\mathrm{ff}}\right)_{\ell, j}^{*}\right] \tag{7.3.41}
\end{equation*}
$$

The time-averaged radiated power crossing the surrounding cylinder over an axial distance $L$ is given by

$$
\begin{equation*}
\mathcal{P}=\int_{-L / 2}^{L / 2} \int_{-\pi}^{\pi}\left(I_{R}\right)_{\mathrm{av}}(R d \theta) d x \tag{7.3.42}
\end{equation*}
$$

Because the $\theta$ and $z$ exponentials in a $\left(P_{\mathrm{ff}}\right)_{m, n}$ term constitute an orthogonal set, the only terms that give a nonzero contribution are those for which $\ell=m$ and $n=j$. This reduces the description to a double sum, which is i.e. $\mathrm{P}=1 / 8$, etc.

$$
\begin{equation*}
\mathcal{P}=\frac{1}{8} \rho_{0} c L \sum_{n=-\infty}^{\infty} \sum_{m=-M}^{M} \frac{\left(k / \kappa_{m}\right)^{2}}{H_{n}^{\prime}\left(\kappa_{m} a\right)^{2}}\left|V_{m, n}\right|^{2} \tag{7.3.43}
\end{equation*}
$$

There is no cross-coupling between individual terms having different axial or circumferential wavenumbers. In other words, each helical surface wave contributes independently to the radiated power. If we had chosen to describe the circumferential
variation of the surface velocity in terms of a real Fourier series, we would have found that the individual $\sin (n \theta)$ and $\cos (n \theta)$ terms do not couple with the radiated power.

The expression for $\mathcal{P}$ is consistent with the earlier theorem regarding radiated power. The region contained between the radiating and surrounding cylinders is not dissipative. Hence, the time-averaged power that flows into the fluid at the surface must flow across the surrounding cylinder. The pressure and radial velocity spread as $1 / R^{1 / 2}$, so the radial intensity is proportional to $1 / R$. The area of the surrounding cylinder over a spatial period $L$ is $2 \pi R L$, so the power that flows across the cylinder over a period does not depend on $R$.

## Axisymmetric Cylindrical Waves

The $n=0$ circumferential harmonic is descriptive of the vibration of a tube that contains a flowing fluid. If the tube is filled with a fluid, and a pressure wave propagates through that fluid, the walls will be pushed outward, thereby inducing an axisymmetric motion of the walls. All points in the cylinder's surface at a specific axial position $z$ share the same radial velocity, so the pressure field is axisymmetric. The following study of the $n=0$ harmonic shows the close relation of the result to the field of an infinite line source.

The surface normal velocity of an axisymmetric wave propagating in the $z$ direction is $V_{s}=V_{0} e^{-i 2 \pi Z / L}$. The corresponding pressure signal may be obtained from Eq.(7.3.38) by letting $m=1$ and setting $V_{0,1}=4 V_{0}$ as the sole nonzero coefficient. (In the case, where the wave propagates in the negative $z$ direction, the $m=-1$ velocity coefficient $V_{0,-1}$ would be the nonzero term.) The farfield pressure obtained by dropping all terms other than $n=0$ in Eq. (7.3.38), then setting $V_{0,1}=4 V_{0}$, is

$$
\begin{equation*}
P=-i \rho_{0} c V_{0}\left(\frac{k a}{\kappa_{1} a}\right) \frac{H_{0}\left(\kappa_{1} R\right)}{H_{0}^{\prime}\left(\kappa_{1} a\right)} e^{-i 2 \pi z / L} \tag{7.3.44}
\end{equation*}
$$

It is instructive to examine the farfield behavior as a function of frequency. There are two independent parameters to consider: $k a$ and $L / a$. Fig. 7.15 is the result of a calculation for four values of $L / a$.


Fig. 7.15 Normalized farfield pressure radiated by a supersonic axisymmetric surface wave propagating axially along an infinite cylinder

The quantity that is plotted is $(R / a)^{1 / 2}\left|P_{\mathrm{ff}}\right| /\left(\rho_{0} c V_{0}\right)$, which is analogous to the manner in which the farfield of a spherical wave is described, except for the difference in the spreading factor. The minimum plotted frequency for each $L / a$ value is $k a=2 \pi a / L$, because no signal radiates to the farfield below that cutoff. At high frequencies, the plotted quantity approaches one. The singularity at $k a=2 \pi a / L$ stems from Eq. (7.3.40) having $\left(\kappa_{1} a\right)^{3 / 2} H_{0}^{\prime}\left(\kappa_{1} a\right)$ in the denominator. If $k a \rightarrow 2 \pi a / L$, then $\kappa_{1} a \rightarrow 0$. Differentiation of the first of Eq.(7.3.13) shows that $H_{0}^{\prime}\left(\kappa_{1} a\right) \rightarrow-2 i /\left(\pi \kappa_{1} a\right)$ for $\kappa_{1} a \ll 1$. This means that the denominator approaches zero as the frequency is reduced to the cutoff value. This singularity means that if we were to induce a vibration amplitude that does not vary with frequency, then a very large pressure would be obtained. The opposite view is that a large radiation impedance indicates that a very large force would be required to sustain the vibration amplitude at a constant value.

The result for infinite $L$ at low frequencies is not singular. The behavior for $L / a=40$ appears to be a transition between the extremes of infinite and small axial wavelengths. An infinite value of $L$ corresponds to a field that is the same at all $z$. Because this field only depends on $R$ and $\theta$, it represents a two-dimensional model. Such models often are used to test concepts and examine physical phenomena. Thus, it is useful to examine this case. We set $2 \pi / L=0$ and $\kappa_{1}=k$ in Eq.(7.3.38), which gives

$$
\begin{equation*}
P=-i \rho_{0} c V_{0} \frac{H_{0}(k R)}{H_{0}^{\prime}(k a)} \tag{7.3.45}
\end{equation*}
$$

There is no axial wave in this case, so there is no minimum frequency for radiation to the farfield. When $H_{0}(k R)$ is replaced in the above with its approximation for large $k R$, and $H_{0}^{\prime}(k a)$ is approximated as $2 i /(\pi k a)$ for small $k a$, we find that the farfield pressure is proportional to $(k a)^{1 / 2}$, rather than the $1 /\left(\kappa_{1} a\right)^{1 / 2}$ proportionality when $L$ is finite and $k a \rightarrow 2 \pi / L$. Thus, in the two-dimensional case, $(R / a)^{1 / 2}|P| /\left(\rho_{0} c V_{0}\right)$ approaches zero as $k a \rightarrow 0$, whereas the finite $L$ case shows a singularity at the lowest $k a$.

We obtained a point source by shrinking a radially vibrating sphere to zero radius. Similarly, we obtain an infinite line source by shrinking the two-dimensional vibrating cylinder to zero radius with the volume velocity per unit length held fixed. The complex amplitude of the radial velocity on the surface is $V_{0}$. If we consider a unit length of the cylinder, the surface area is $2 \pi a$, so the volume velocity per unit length is $\hat{Q}_{s}=2 \pi a V_{0}$. We use this definition to replace $V_{0}$, and also invoke the small $k a$ approximation of $H_{0}^{\prime}(k a)$, see Eq. (7.3.13). Doing so converts the two-dimensional solution to the line source pressure,

$$
\begin{equation*}
P=\frac{1}{4} \rho_{0} \omega \hat{Q}_{s} H_{0}(k R) \tag{7.3.46}
\end{equation*}
$$

Recall that the free-space Green's function was defined in Eq. (6.4.29) to be the solution of the Helmholtz equation with a Dirac delta inhomogeneous term, with the additional specification that it has unit mass acceleration. The mass acceleration per
unit length of the line source is $i \omega \rho_{0} \hat{Q}_{s}$, so the two-dimensional free-space Green's function is

$$
\begin{gather*}
G\left(\bar{x}, \bar{x}_{s}\right)=\frac{1}{4 i} H_{0}(k R)  \tag{7.3.47}\\
R=\left|\bar{x}-\bar{x}_{s}\right|=\left[\left(x-x_{s}\right)^{2}+\left(y-y_{s}\right)^{2}\right]^{1 / 2}
\end{gather*}
$$

It often is more convenient to use the monopole amplitude of a line source, rather than the volume velocity. The monopole amplitude $A$ for a point source was defined to be the coefficient of the position-dependent terms. We adopt the same definition here, so that the line source pressure may be written in either of two forms,

$$
\begin{gather*}
P=i \omega \rho_{0} \hat{Q}_{s} G\left(\bar{x}, \bar{x}_{s}\right)=A H_{0}(k R)  \tag{7.3.48}\\
\hat{Q}_{s}=2 \pi a V_{0}, \quad A=\frac{1}{4} \omega \rho_{0} \hat{Q}_{s}
\end{gather*}
$$

This development is important because concepts and procedures that feature Green's functions for three-dimensional situations are equally applicable for a twodimensional model. For example, we may employ the method of images to analyze the field radiated by a line source above a rigid or pressure-release plane.

EXAMPLE 7.7 The surface of a cylinder whose length is $L$ undergoes a uniform radial vibration, such that $v_{R}=\operatorname{Re}\left(v_{0} e^{i \omega t}\right)$ everywhere. The cylinder has finite length $L$, and its radius is very small relative to the acoustic wavelength $2 \pi / k$. The origin of the $x y z$ coordinate system is situated at the center of the cylinder, with the $z$-axis coincident with the cylinder's centerline. Derive an expression for the pressure at an arbitrary field point. Use this expression to evaluate $|P|$ as a function of distance from the centerline along the $x$-axis. Consider $k L=0.5,5$, and 50 , and compare each result to the field of an infinite line source.

## Significance

This is one of the few problems featuring a finite length cylinder that may be solved without recourse to advanced numerical techniques. In addition to providing a quantitative picture that suggests when it might be acceptable to take a cylinder's length to be infinite, it is a useful reminder of the technique for superposing point source fields.

## Solution

The cylinder reduces to a line source when its radius is much less than a wavelength. The case of an infinite length is described by Eq.(7.3.48), which may be applied directly. To evaluate the finite length cases, we return to the concept of a source distribution. The cylinder at zero radius may be considered to be a continuous
distribution of point sources that are differential segments of the line. A sketch of the system shows the radial distance $r$ from the cylinder segment to the field point, whereas $R$ is the perpendicular distance from the field point to the cylinder. The distance from the origin to a differential segment is denoted as $s$ in order to distinguish this distance from the coordinate $z$ of the field point. Correspondingly, $d s$ is the length of the segment. The field is axisymmetric, so the field point may be situated in the $x z$ plane, as it is shown in Fig. 1.


Figure 1.

The radial velocity of a differential segment is $V_{R}=v_{0}$ and its surface area is $2 \pi a d s$, so the complex amplitude of its volume velocity is $d \hat{Q}=\hat{Q}_{s} d s=$ $v_{0}(2 \pi a d s)$. We retain $\hat{Q}_{s}$ rather than $a v_{0}$ because $v_{0}$ is not meaningful in the limit as $a \rightarrow 0$. The corresponding mass acceleration of the element is $i \omega \rho_{0} \hat{Q}_{s} d s$. The radial distance from the differential element to the field point is $r=\left[R^{2}+(s-z)^{2}\right]^{1 / 2}$, and the (three-dimensional) Green's function is $G\left(\bar{x}, \bar{x}_{s}\right)=\exp (-i k r) / 4 \pi r$. The pressure at the field point is the sum of the signal received from each differential segment. In other words, it is an integral,

$$
\begin{align*}
P(R, z) & =\int G\left(\bar{x}, \bar{x}_{s}\right)\left(i \omega \rho_{0} \hat{Q}_{s} d s\right) \\
& =i \omega \rho_{0} \hat{Q}_{s} \int_{-L / 2}^{L / 2} \frac{e^{-i k\left[R^{2}+(s-z)^{2}\right]^{1 / 2}}}{4 \pi\left[R^{2}+(s-z)^{2}\right]^{1 / 2}}\left(i \omega \rho_{0}\right)\left(2 \pi a v_{0}\right) d s \tag{1}
\end{align*}
$$

The requested evaluation places the field point on the $x$-axis. Setting $z=0$ in the general expression gives

$$
\begin{equation*}
P(R, 0)=i \omega \rho_{0} \frac{\hat{Q}_{s}}{2 \pi} \int_{0}^{L / 2} \frac{e^{-i k\left(R^{2}+s^{2}\right)^{1 / 2}}}{\left(R^{2}+s^{2}\right)^{1 / 2}} d s \tag{2}
\end{equation*}
$$

where the fact that the integrand only depends on $s^{2}$ allows us to double the integral over positive $s$.

The integral in Eq. (2) is not one that can be found in a standard tabulation. Nevertheless, let us consider an analytical integration. The square root is a complicating feature, so it is reasonable to consider changing the integration variable. We can transform from $s$ to the angle $\psi$ in the sketch. The trigonometric relations are $s=R \tan \psi$, $\left(R^{2}+s^{2}\right)^{1 / 2}=R \sec \psi$, and $d s=R d \psi(\sec \psi)^{2}$. This does not make the integrand more amenable to analysis. Another transformation that eliminates the square root
is $s=R \sinh \zeta$, for which $\left(R^{2}+s^{2}\right)^{1 / 2}=R \cosh \zeta$ and $d s=R(\cosh \zeta) d \zeta$. This is promising because it converts the integral to

$$
\begin{equation*}
P(R, 0)=i \omega \rho_{0} \frac{\hat{Q}_{s}}{2 \pi} \int_{0}^{\sinh ^{-1}\left(\frac{L}{2 R}\right)} e^{-i k R \cosh \zeta} d \zeta \tag{3}
\end{equation*}
$$

The complex exponential may be decomposed into its real and imaginary parts. A search through integral tables yields results for the case where $L$ is infinite, but not finite. It is worth the effort to consider that limit, because it should agree with Eq. (7.3.48), and thereby confirm our analysis. The tabulated formulas ${ }^{14}$ are

$$
\begin{aligned}
J_{0}(k R) & =\frac{2}{\pi} \int_{0}^{\infty} \sin (k R \cosh \zeta) d \zeta \\
N_{0}(k R) & =-\frac{2}{\pi} \int_{0}^{\infty} \cos (k R \cosh \zeta) d \zeta
\end{aligned}
$$

Thus, if $L$ were infinite, Eq. (3) would be

$$
\begin{aligned}
P(R, 0) & =i \omega \rho_{0} \frac{Q_{0}}{2 \pi} \int_{0}^{\infty}[\cos (k R \cosh \zeta)-i \sin (k R \cosh \zeta)] d \zeta \\
& =i \omega \rho_{0} \frac{Q_{0}}{2 \pi}\left(\frac{\pi}{2}\right)\left[-N_{0}(k R)-i J_{0}(k r)\right]
\end{aligned}
$$

The term in the bracket is $-i H_{0}(k R)$, so the pressure for the case of infinite $L$ reduces to

$$
\begin{equation*}
P_{\mathrm{inf}}(R, 0)=i \omega \rho_{0} \hat{Q}_{s}\left(\frac{1}{4 i} H_{0}(k R)\right) \tag{4}
\end{equation*}
$$

This is the same as (7.3.48). In addition to confirming the analysis, this result tells us that evaluating Eq. (3) for increasing values of $L / R$ should yield results that are close to those obtained for an infinite length.

The only recourse for a finite value of $L$ is a numerical evaluation of the integral. The numerical integration scheme we employ could be one of our own construction using a technique like Simpson's rule. Instead, we shall use a routine provided by MATLAB, which is quadl. Although the integrand in Eq. (2) does not have a singularity, there is a potential difficulty in evaluating it for large values of $k R$. The cause of this difficulty is that $k R$ is like the frequency of harmonic terms. A large value of $k R$ leads to an integrand that is a rapidly oscillating function of $\zeta$. Some say that the integrand has a singularity at infinite $\zeta$. This property must be acknowledged in the numerical integration, either by using an adaptive routine, or by explicitly increasing the number of integration points as the value of $k R$ increases.

The maximum value of $k R$ has not been specified. We know that if $R \gg L$, then the three-dimensional farfield approximation should apply. Because $R / L \equiv$

[^11]$k R /(k L)$, this necessitates adjusting the range of $k R$ commensurate with $k L$. Thus, the computation first sets the $k L$ value. A program loop evaluates Eq. (2) at each $k R$ value from zero to $60 k L$, with an increment of $k L / 2$.

Figures 2-4 display the results for each $k L$. The data indicates that the finite and infinite line sources for $k L \geq 4$ agree well in the region from $R=0$ out to the distance at which $k R=O(k L)$, in other words, $0 \leq R \leq O(L)$. Beyond that range, the infinite line source model overpredicts the pressure. For the shortest length, $k L=0.5$, the pressure for an infinite line source everywhere exceeds the pressure radiated by the finite cylinder.


Figure 2.


Figure 3.


Figure 4.

These trends have a simple explanation. If the length is comparable to, or larger than, one wavelength, a field point that is not too distant from the cylinder will most strongly be affected by the signals that radiate from the closest source points $(z \approx 0)$. As the field point is moved farther outward, the contribution of the sources near the ends becomes relatively more significant because all points on the cylinder approach being equidistant from the field point. In the case where $L$ is much smaller than a wavelength, the cylinder is acoustically compact. The best simple model in that case is a point source, rather than an infinite line source.

Numerical simulations and experimental measurements of a variety of systems have found that these trends are general. This knowledge allows us to use infinite length models to investigate the nearfield of a finite length body, provided that the length of that body is not much less than a wavelength. A useful corollary is that the radiation impedance of a finite length cylinder may be expected to be close to that of an infinite cylinder if the length is comparable to, or larger than, $2 \pi / k$. We will exploit this attribute in Exercise 7.10.

## Field Generated by a Vibrating Cable or Beam

When a flexural wave propagates along the $z$-axis of a cable or beam, all points on a specific cross section displace perpendicularly to the axis by the same amount. In other words, each cross section translates without alteration of its shape. We shall define the $x$-axis to be such that the displacement $\bar{u}$ lies in the $x z$ plane, so that

$$
\begin{equation*}
\bar{u}=\operatorname{Re}\left(\frac{V_{x}}{i \omega} e^{-i 2 \pi z / L} e^{i \omega t}\right) \bar{e}_{x} \tag{7.3.49}
\end{equation*}
$$

The component of this displacement normal to a circular cross section is $\bar{u} \cdot \bar{e}_{R}$. We measure the circumferential angle $\theta$ relative to the $x$-axis, so $\bar{e}_{x} \cdot \bar{e}_{R}=\cos \theta$. Correspondingly, the surface velocity is

$$
\begin{equation*}
V_{s}=V_{x} e^{-i 2 \pi z / L} \cos \theta \equiv \frac{1}{2} V_{x} e^{-i 2 \pi z / L}\left(e^{-i \theta}+e^{i \theta}\right) \tag{7.3.50}
\end{equation*}
$$

This representation tells us that the surface wave consists of $n=1$ and $n=-1$ circumferential harmonics. A comparison with Eq. (7.3.36) shows that the only nonzero coefficients in this motion are $V_{1,(-1)}=V_{1,1}=2 V_{x}$. Thus, the pressure generated by a supersonic wave propagating along a cable or beam is

$$
\begin{equation*}
P=-i \rho_{0} c V_{x} \frac{k H_{1}\left(\kappa_{1} R\right)}{\kappa_{1} H_{1}^{\prime}\left(\kappa_{1} a\right)} e^{-i 2 \pi z / L} \cos \theta \tag{7.3.51}
\end{equation*}
$$

Figure 7.16 depicts the frequency dependence of the farfield pressure amplitude normalized by $\rho_{0} c V_{0}(R / a)^{1 / 2}$. The minimum frequency for radiation is $k a=$ $2 \pi a / L$. The high-frequency limit is seen to be a value of one, as it was for $n=0$. Rather than being singular, the transition from a subsonic to supersonic surface wave is marked by a zero value for the pressure. To see why this is so, we observe that $H_{1}\left(\kappa_{1} R\right)$ is proportional to $1 /\left(\kappa_{1} R\right)^{1 / 2}$ at large $\kappa_{1} R$ (farfield), while $H_{1}^{\prime}\left(\kappa_{1} a\right) \rightarrow-(2 i / \pi) /\left(\kappa_{1} a\right)^{2}$ as $\kappa_{1} a \rightarrow 0$ (small radius). The pressure that results from combining these limiting forms is

$$
\begin{equation*}
P=\rho_{0} c V_{x}\left(\frac{\pi}{2}\right)^{1 / 2} k a\left(\kappa_{1} a\right)^{1 / 2}\left(\frac{a}{R}\right)^{1 / 2} e^{-i\left(\kappa_{1} R-3 \pi / 4-2 \pi z / L\right)} \cos \theta \tag{7.3.52}
\end{equation*}
$$

Thus rather than $P$ being singular, it approaches a proportionality to $\left(\kappa_{1} a\right)^{1 / 2}$ as $k \rightarrow 2 \pi / L$. However, if $L / a$ is not too large, the figure indicates that values of $k a$ slightly larger than $2 \pi a / L$ give rise to large pressures.


Fig. 7.16 Normalized farfield pressure radiated by a supersonic flexural wave propagating axially along an infinite cylinder

A two-dimensional model of a translating beam is obtained by letting $L \rightarrow \infty$. When we studied point sources, we found that a translating sphere generates a dipole field, and that a dipole field may be obtained from a Taylor series expansion of the field generated by two oppositely phased monopoles. Let us see whether the same is true in the two-dimensional case. We consider two line sources at $x=d / 2$ and $x=-d / 2$, where their placement in the $x z$ plane gives symmetry relative to that plane, which is the property of $\cos \theta$. The mass accelerations are $\pm i \omega \rho_{0} \hat{Q}_{s}$, so that

$$
\begin{equation*}
P_{\text {dipole }}=i \omega \rho_{0} \hat{Q}_{s} G\left(\bar{x}, \bar{x}_{0}+\frac{d}{2} \bar{e}_{x}\right)-i \omega \rho_{0} \hat{Q}_{s} G\left(\bar{x}, \bar{x}_{0}-\frac{d}{2} \bar{e}_{x}\right) \tag{7.3.53}
\end{equation*}
$$

where the Green's function is as given in Eq. (7.3.47) and $\bar{x}_{0}$ is situated on the $z$-axis. Furthermore, for a two-dimensional analysis, we take $z=0$ for both the field and source points. We apply a Taylor series to both terms, and switch the gradient at $\bar{x}_{0}$ for the gradient at $\bar{x}$, which changes the sign associated with each gradient. Thus, we have

$$
\begin{equation*}
P=i \omega \rho_{0} \hat{Q}_{s}\left[-d \bar{e}_{x} \cdot \nabla G\left(\bar{x}, \bar{x}_{0}\right)\right] \tag{7.3.54}
\end{equation*}
$$

Now we observe that $G\left(\bar{x}, \bar{x}_{0}\right)$ depends only on the distance $R$ between the points in the $x y$ plane, and that the gradient of $R$ at the field point is a unit vector pointing in the direction of increasing $R$, which is $\bar{e}_{R}$. This leads to

$$
\begin{align*}
P & =-i \omega \rho_{0} \hat{Q}_{s} d\left(\bar{e}_{x} \cdot \bar{e}_{R}\right) \frac{\partial}{\partial R} G\left(\bar{x}, \bar{x}_{0}\right)  \tag{7.3.55}\\
& =-\frac{1}{4} \omega \rho_{0} \hat{Q}_{s} d(\cos \theta) H_{1}(k R)
\end{align*}
$$

The dipole moment $D$ is the product of the monopole amplitude and the separation distance. Introduction of $A$ from Eq. (7.3.48) into the preceding leads to

$$
\begin{equation*}
P=D H_{1}(k R) \cos \theta, \quad D=A d=-\frac{1}{4} \omega \rho_{0} \hat{Q}_{s} d \tag{7.3.56}
\end{equation*}
$$

A comparison of this result with Eq. (7.3.51) for $a / L=0$ shows that the dipole moment associated with a translating rigid cylinder is

$$
\begin{equation*}
D_{\mathrm{cyl}}=-\frac{i \rho_{0} c V_{x}}{H_{1}^{\prime}\left(\kappa_{1} a\right)} \tag{7.3.57}
\end{equation*}
$$

In principle, we could mirror every development using three-dimensional point sources with an analogous one for the two-dimensional field for a set of infinite line source. For example, we could develop a multipole expansion in two dimensions. In practice, this is seldom done because of the artificiality of the infinite length model.

EXAMPLE 7.8 When the string of a musical instrument is plucked, the free vibration that ensues is a standing wave that consists of a series of mode functions. The displacement of mode number $m$ is proportional to $\sin (m \pi z / \ell) \cos \theta$ and its frequency is $m \omega_{1}$, where $\omega_{1}$ is the fundamental frequency and $\ell$ is the string's length. In the situation of interest, the displacement amplitude $U_{m}$ of modes 1 to 10 is known to be inversely proportional to the square of the mode number, and the contributions of higher modes are negligible. Thus, $U_{m}=U_{1} / m^{2}$ if $m=1, \ldots, 10, U_{m}=0$ if $m>10$, where $U_{1}$ is the amplitude of the fundamental. Consider a model of a cello string that ignores the presence of the cello body and takes the string to be an infinitely cylinder that undergoes an axially periodic vibration. The string's length is 670 mm and its diameter is 1.20 mm . For the case where the sound pressure level at 2 m from the string is measured to be 75 dB and the fundamental frequency is 440 Hz , estimate the RMS displacement amplitude of the string.

## Significance

Inverse radiation problems, in which a system's vibration is deduced from measured properties of the pressure field, help reinforce understanding of an analytical solution. They also mimic, in a small way, the nature of some experimental investigations.

## Solution

We know the pressure at a field point and must deduce properties of the vibration that generated it. Our strategy is to reverse the problem. The only quantity not specified for the displacement is the amplitude $U_{1}$ of the $m=1$ mode. It will be determined by computing $\left(p^{2}\right)_{\text {av }}$ in terms of $U_{1}$, then matching that result to the value corresponding to $75 \mathrm{~dB} / / 20 \mu \mathrm{~Pa}$, which is $0.01265 \mathrm{~Pa}^{2}$.

The specified displacement of the cable, which we take to be in the $x$ direction, is

$$
\begin{equation*}
\bar{u}=\operatorname{Re} \sum_{m=1}^{10} \frac{U_{1}}{m^{2}} \sin \left(\frac{m \pi z}{\ell}\right) e^{i m \omega_{1} t} \bar{e}_{x} \tag{1}
\end{equation*}
$$

We differentiate this expression with respect to time, then take the radial component in order to determine the normal velocity on the surface. Because $\bar{e}_{x} \cdot \bar{e}_{r}=\cos \theta$, the resulting expression is

$$
v_{s}=\operatorname{Re} \sum_{m=1}^{10} \frac{i \omega_{1} U_{1}}{m} \sin \left(\frac{m \pi z}{\ell}\right)(\cos \theta) e^{i m \omega_{1} t}
$$

We will use Eq. (7.3.51) to describe the pressure, so we employ Euler's identity to decompose $v_{s}$ into waves traveling in opposite directions, which gives

$$
\begin{equation*}
v_{s}=\operatorname{Re} \sum_{m=1}^{10} \frac{U_{1} \omega_{1}}{4 m}\left(e^{i m \pi z / \ell}-e^{-i m \pi z / \ell}\right)\left(e^{i \theta}+e^{-i \theta}\right) e^{i m \omega_{1} t} \tag{2}
\end{equation*}
$$

This expression tells us that each time harmonic consists of a pair of helical waves at $n= \pm 1$ that spiral about the $z$-axis in opposite senses. The velocity coefficients are $V_{m, 1}=V_{m,-1}=-V_{-m, 1}=-V_{-m,-1}=\omega U_{1} /(2 m)$.

The nature of the pressure generated by each wave depends on whether it is supersonic. The frequency of a term in Eq. (2) is $m \omega_{1}$ and its axial wavenumber is $m \pi / \ell$, so that

$$
\begin{equation*}
\kappa_{m}=\left[\left(\frac{m \omega_{1}}{c}\right)^{2}-\left(\frac{m \pi}{\ell}\right)^{2}\right]^{1 / 2}=m \kappa_{1}=6.780 m \mathrm{~m}^{-1} \tag{3}
\end{equation*}
$$

All values of $\kappa_{m}$ are real, which means that all of the terms in Eq. (1) are supersonic helical waves.

Each term in Eq. (2) occurs at a different frequency, so the time factor must be retained when Eq. (7.3.51) is used to evaluate the pressure. In each term of that expression, we replace $k$ with $m \omega_{1} / c, \kappa_{1}$ with $m \kappa_{1}$, and $2 / L$ with $m / \ell$. The result is

$$
\begin{align*}
p & =\operatorname{Re}\left[-i \rho_{0} c \sum_{m=1}^{10} \frac{U_{1} \omega_{1}}{2 m}\left(e^{i m \pi z / \ell}-e^{-i m \pi z / \ell}\right)\left(\frac{\omega_{1}}{c \kappa_{1}}\right) \frac{H_{1}\left(m \kappa_{1} R\right)}{H_{1}^{\prime}\left(m \kappa_{1} a\right)}(\cos \theta) e^{i m \omega_{1} t}\right] \\
& =\operatorname{Re}\left[-i \rho_{0} \frac{\omega_{1}^{2}}{\kappa_{1}} U_{1} \sum_{m=1}^{10} \frac{H_{1}\left(m \kappa_{1} R\right)}{m H_{1}^{\prime}\left(m \kappa_{1} a\right)} \sin \left(\frac{m \pi z}{\ell}\right)(\cos \theta) e^{i m \omega_{1} t}\right] \tag{4}
\end{align*}
$$

Each term in Eq. (4) represents an oscillation at a distinct frequency, so the meansquared pressure is a sum of squares of the individual amplitudes. Thus, the predicted mean-square pressure is

$$
\begin{equation*}
\left(p^{2}\right)_{\mathrm{av}}=\frac{1}{2} \sum_{m=1}^{\infty} P_{m} P_{m}^{*}=\frac{1}{2}\left(\frac{\rho_{0} \omega_{1}^{2}}{\kappa_{1}}\left|U_{1}\right|\right)^{2} \sum_{m=1}^{10}\left|\frac{H_{1}\left(m \kappa_{1} R\right)}{m H_{1}^{\prime}\left(m \kappa_{1} a\right)}\right|^{2}\left[\sin \left(\frac{m \pi z}{\ell}\right)\right]^{2}(\cos \theta)^{2} \tag{5}
\end{equation*}
$$

The transverse distance for the pressure measurement is specified to be $R=2 \mathrm{~m}$, but the values of $z$ and $\theta$ at which the pressure was measured are not given. It is reasonable to assume that the value of $\theta$ is that which leads to a maximum, which is
$\theta=0$ or $\pi$, both of which give $(\cos \theta)^{2}=1$. In contrast, no value of $z$ is evident as leading to a maximum. For example, if we select the midpoint, $z=\ell / 2$, the terms for even $m$ will be zero. Thus, let use the value of $\left(p^{2}\right)_{\mathrm{av}}$ averaged over the length of the string, which is

$$
\begin{equation*}
\left(p^{2}\right)_{\mathrm{av}, \mathrm{av}}=\left.\frac{1}{\ell} \int_{0}^{\ell}\left(p^{2}\right)_{\mathrm{av}, \mathrm{av}}\right|_{\theta=0} d z=\frac{1}{4}\left(\frac{\rho_{0} \omega_{1}^{2}}{\kappa_{1}}\left|U_{1}\right|\right)^{2} \sum_{m=1}^{10}\left|\frac{H_{1}\left(m \kappa_{1} R\right)}{m H_{1}^{\prime}\left(m \kappa_{1} a\right)}\right|^{2} \tag{6}
\end{equation*}
$$

The result of substituting all quantities into Eq. (6) is $\left(p^{2}\right)_{\mathrm{av}, \mathrm{av}}=784.3\left|U_{1}\right|^{2}$. Matching this to the measured mean-squared pressure of $0.01265 \mathrm{~Pa}^{2}$ leads to $\left|U_{1}\right|=$ 4.016 mm . With this, we have fully characterized the displacement. Its mean-squared value is obtained by averaging $\bar{u} \cdot \bar{u}$ in Eq. (1) over $0<t<\pi / \omega_{1}$ and $0<x<\ell$. The result is

$$
\begin{equation*}
\left(u_{x}^{2}\right)_{\mathrm{av}}=\frac{\left|U_{1}\right|^{2}}{4} \sum_{m=1}^{10} \frac{1}{m^{4}}=4.363\left(10^{-6}\right) \mathrm{m}^{2} \Longrightarrow\left(u_{x}\right)_{\mathrm{rms}}=2.08 \mathrm{~mm} \tag{7}
\end{equation*}
$$

The actual displacement amplitude of the strings of a cello is barely perceptible to the eye. The value we have determined is too large. The discrepancy between our analysis and reality is only partially the consequence of taking the string's vibration to be spatially periodic and using a model of an infinitely long string. The primary fault is that the strings are an extremely minor contributor to the sound emitted by a stringed musical instrument. The role of the strings is to induce vibration of the instrument's body. This vibration radiates sound directly, and it also excites the cavity enclosed by the body. The sound created in the cavity is emitted through the open ports of the body.

### 7.3.5 Finite Length Effects

Two aspects of the vibrating cylinder model we have created limit its practical utility. The most obvious one is taking the cylinder to be infinitely long. Unfortunately, there are no analytical solutions for radiation of finite length cylinders. If one requires an accurate solution of such problems, all available tools use numerical techniques. We will examine a few later in this chapter.

Even within the context of a model that takes a cylinder to have infinite length, there is another unrealistic aspect. We have considered helical waves and axially periodic vibrations, which leads to the question of how would one generate such a disturbance? A more realistic model allows the surface velocity to depend in an arbitrary manner on the axial distance, subject to the condition that it be negligible at large distances in either direction. In order to not obscure the analysis, we will restrict attention to the case of a single circumferential harmonic. The solutions may be superposed to create the Fourier series for an arbitrary circumferential excitation.

The surface velocity for this investigation is

$$
\begin{equation*}
v_{s}=\operatorname{Re}\left(f(z) e^{-i n \theta} e^{i \omega t}\right), \quad \lim _{|z| \rightarrow \infty}|f(z)| \rightarrow 0 \tag{7.3.58}
\end{equation*}
$$

This distribution is not periodic in the axial direction, which is equivalent to saying that its period is infinite. The conditions imposed on $f(z)$ fit the requirements for its Fourier transform to exist. The transform, which is denoted as $\tilde{F}(\mu)$, is

$$
\begin{equation*}
\tilde{F}(\mu)=\int_{-\infty}^{\infty} f(z) e^{+i \mu z} d z \tag{7.3.59}
\end{equation*}
$$

As explained in Appendix B, the Fourier transform of $f(z)$ essentially describes a Fourier series whose wavenumber is a continuous spectrum $\mu$, rather than a set of discrete values $n \mu_{1}$. The series coefficient at each $\mu$ is $\tilde{F}(\mu) d \mu$. This perspective is especially evident when the inverse Fourier transform is used to synthesize $f(z)$. It tells us that

$$
\begin{equation*}
V_{s}=\frac{1}{2 \pi} \int_{-\infty}^{\infty} \tilde{F}(\mu) e^{-i(\mu z+n \theta)} d \mu \tag{7.3.60}
\end{equation*}
$$

The elementary view of an integral is that it is an infinite sum of terms described by its integrand. From that viewpoint, the above description states that the surface velocity is an infinite sum of helical waves whose axial wavenumber is $\mu$, whose circumferential wavenumber is $n$, and whose complex amplitude is $\tilde{V}(\mu) d \mu$. The fact that the number of terms in the sum is infinite does not affect the validity of superposing the pressure field for each helical wave. Equation (7.3.21) gives the pressure for a single helical wave, so the superposition gives

$$
\begin{equation*}
P(R, \theta, z)=-i \rho_{0} c \int_{-\infty}^{\infty} \tilde{F}(\mu) \frac{k H_{n}(\kappa(\mu) R)}{\kappa(\mu) H_{n}^{\prime}(\kappa(\mu) a)} e^{-i(\mu z+n \theta)} d \mu \tag{7.3.61}
\end{equation*}
$$

where

$$
\begin{equation*}
\kappa(\mu)=\left(k^{2}-\mu^{2}\right)^{1 / 2} \tag{7.3.62}
\end{equation*}
$$

The transverse wavenumber has been written as $\kappa(\mu)$ as a reminder that its dependence must be recognized when the integral is evaluated. The portion of the integrand in Eq. (7.3.61) that is a factor of $e^{-i \mu z}$ is a Fourier transform. Thus, we have obtained a representation of $P$ as an inverse Fourier transform.

The preceding is a heuristic derivation. A more rigorous analysis provides understanding as to how one may apply the Fourier transform in other situations. We begin with the argument that the dependence of $P$ on the axial position also may be represented by a Fourier transform. In addition to depending on the axial wavenumber $\mu$, this transform may depend on the transverse distance $R$, whereas $\exp (-i n \theta)$ must describe the circumferential dependence. If we knew the transformed function, which is denoted as $\tilde{P}(R, \mu)$, then the inverse Fourier transform would enable us to recreate $P(R, \theta . z)$ according to

$$
\begin{equation*}
P(R, \theta, z)=\frac{1}{2 \pi} \int_{-\infty}^{\infty} \tilde{P}(R, \mu) e^{-i \mu z} e^{-i n \theta} d \mu \tag{7.3.63}
\end{equation*}
$$

This ansatz must satisfy the Helmholtz equation in cylindrical coordinates. Spatial derivatives of the above are

$$
\begin{align*}
\frac{\partial}{\partial R} P(R, \theta, z) & =\frac{1}{2 \pi} \int_{-\infty}^{\infty} \frac{\partial}{\partial R} \tilde{P}(R, \mu) e^{-i \mu z} e^{-i n \theta} d \mu \\
\frac{\partial}{\partial \theta} P(R, \theta, z) & =\frac{1}{2 \pi} \int_{-\infty}^{\infty}(-i n) \tilde{P}(R, \mu) e^{-i \mu z} e^{-i n \theta} d \mu  \tag{7.3.64}\\
\frac{\partial}{\partial z} P(R, \theta, z) & =\frac{1}{2 \pi} \int_{-\infty}^{\infty}(-i \mu) \tilde{P}(R, \mu) e^{-i \mu z} e^{-i n \theta} d \mu
\end{align*}
$$

Second-order derivatives follow directly from these forms, so substitution of the assumed solution into the Helmholtz equation leads to

$$
\begin{equation*}
\int_{-\infty}^{\infty}\left[\frac{d^{2} \tilde{P}}{d R^{2}}+\frac{1}{R} \frac{d \tilde{P}}{d R}+\left(\kappa(\mu)^{2}-\frac{n^{2}}{R^{2}}\right) \tilde{P}\right] e^{-i \mu z} e^{-i n \theta} d \mu=0 \tag{7.3.65}
\end{equation*}
$$

The preceding must be satisfied at all $z$, which will only be true if the bracketed term in the integrand vanishes. Thus, $\tilde{P}(R, \mu)$ must be a solution of Bessel's equation. As previous, an outgoing wave is obtained by taking it to be a Hankel function of the second kind. It may be multiplied by a coefficient that depends on $\mu$ and $n$, so we now have

$$
\begin{equation*}
P(R, \theta, z)=\int_{-\infty}^{\infty} B(\mu, n) H_{n}(\kappa(\mu) R) e^{-i \mu z} e^{-i n \theta} d \mu \tag{7.3.66}
\end{equation*}
$$

The coefficient $B(\mu, n)$ is set by the boundary condition, which matches $V_{s}$ to the complex particle velocity in the radial direction at $R=a$. A derivative with respect to $R$ for Euler's equation operates on the Hankel function, so using Eq. (7.3.60) to represent the surface velocity leads to

$$
\begin{equation*}
\int_{-\infty}^{\infty} \kappa(\mu) B(\mu, n) H_{n}^{\prime}(\kappa(\mu) a) e^{-i \mu z} e^{-i n \theta} d \mu=-i \omega \rho_{0} \int_{-\infty}^{\infty} \tilde{F}(\mu) e^{-i(\mu z+n \theta)} d \mu \tag{7.3.67}
\end{equation*}
$$

Here too, the equality must hold for all $z$, so the integrands must be equal. From this, we solve for $B(\mu, n)$. Backsubstituting that results into the integral representation of $P$ leads to Eq. (7.3.61), which verifies the heuristic approach.

The spectrum of surface waves consists of supersonic waves for $|\mu|<k$ and subsonic waves for $|\mu|>k$. For subsonic waves, let us set $\kappa(\mu)=-i \beta(\mu)$, where $\beta=\left(\mu^{2}-k^{2}\right)^{1 / 2}$. It is helpful for evaluations to rewrite Eq. (7.3.61) in a form that splits the two in order to explicitly display the contributions of the supersonic and
subsonic spectra. For this, we note that neither $\kappa$ nor $\beta$ depends on the sign of $\mu$, so we may reduce the integral to extend only over $\mu \geq 0$. These operations lead to

$$
\begin{align*}
P(R, \theta, z)= & -i \rho_{0} c \int_{0}^{k}[\tilde{F}(\mu)+\tilde{F}(-\mu)] \frac{k H_{n}(\kappa(\mu) R)}{\kappa(\mu) H_{n}^{\prime}(\kappa(\mu) a)} e^{-i(\mu z+n \theta)} d \mu \\
& +\rho_{0} c \int_{k}^{\infty}[\tilde{F}(\mu)+\tilde{F}(-\mu)] \frac{k K_{n}(\beta(\mu) R)}{\beta(\mu) K_{n}^{\prime}(\beta(\mu) a)} e^{-(\mu z+n \theta)} d \mu \tag{7.3.68}
\end{align*}
$$

An important aspect of this representation is that the contribution from $\mu>k$ decays rapidly with increasing $R$ and $z$. Therefore, the second integral may be ignored if we seek a farfield description.

Equation (7.3.68) is an inverse Fourier transform. It is an implementation of a general analytical approach known as an angular spectrum decomposition. We will encounter in the next chapter another system that has been analyzed by this technique. Many research papers have derived solutions in the form of inverse Fourier transforms, and then used sophisticated tools of mathematical analysis to identify important properties. Such analysis exceeds the scope of this book. An alternative is to employ numerical methods. The last part of Appendix B describes how to use an FFT numerical procedure to evaluate the Fourier transform of a spatial function, and how to convert a Fourier transform back to a spatial function by application of an inverse FFT procedure. Thus, it might seem viable to employ FFT techniques to evaluate Eq. (7.3.61). Despite the apparent simplicity of this approach, there are several issues regarding aliasing and wraparound error, which were discussed in Chap. 1, that must be addressed. An additional complication arises as either $R$ or $z$ is increased, because the integrands become rapidly oscillating functions of $\mu$. This raises the sampling requirements, and sometimes leads to numerical instabilities.

Ultimately, evaluation of the pressure field would have limited use. As was shown in Example 7.7, the fact that the cylinder is not infinite becomes important at radial distances that exceed the length. Furthermore, even if we are only interested in the field close to a cylinder, end effects cannot be ignored at locations that are close to an end.

### 7.4 Kirchhoff-Helmholtz Integral Theorem

It should be apparent at this juncture that we need tools other than separation of variables solutions of the Helmholtz equation. The first step toward that objective is application of the definition of a Green's function to derive the Kirchhoff-Helmholtz integral theorem, which we will refer to as the KHIT. It is the basis for some common numerical modeling technique, but it is not a solution of any problem. Rather, it will tell us how to evaluate the pressure at a field point when we know the pressure
and normal velocity on the surface from which this signal emanates. The laws of mechanics, as well as our experience thus far, tell us that both quantities cannot be specified at the same location. (This fact is merely an extension of Newton's second law for a particle, which tells us that we can determine the velocity as a function of time if we know the force history, or we can determine the force required to attain a specified velocity history, but we cannot specify both histories.) The KHIT version we will derive pertains to the response in the frequency domain.

### 7.4.1 Derivation for an Acoustic Cavity

The development begins with the simultaneous consideration of two acoustic fields: the frequency domain pressure radiated by vibration of the boundary and a Green's function. Usually, the free-space Green's function is the one we would use, but all steps will be equally valid of the Green's function satisfies boundary conditions for the system at hand. We let $\mathcal{V}$ denote the domain of interest, and $\mathcal{S}$ is its boundary. For our initial effort, we consider $\mathcal{V}$ to be enclosed by $\mathcal{S}$. In other words, it is a cavity whose shape is arbitrary. The configuration is depicted in Fig. 7.17.

Fig. 7.17 Configuration of an acoustic cavity


The pressure at a field point $\bar{x}$ satisfies the Helmholtz equation. The Green's function for this point due to a source at $\bar{x}_{0}$, which is $G\left(\bar{x}_{0}, \bar{x}\right)$, satisfies an inhomogeneous version of that equation, specifically

$$
\begin{equation*}
\nabla^{2} P+k^{2} P=0, \quad \nabla^{2} G+k^{2} G=-\delta\left(\bar{x}-\bar{x}_{0}\right) \tag{7.4.1}
\end{equation*}
$$

We multiply the equation for $P$ by $G\left(\bar{x}_{0}, \bar{x}\right)$ and the equation for $G$ by $P(\bar{x})$. Taking the difference yields

$$
\begin{equation*}
G\left(\bar{x}_{0}, \bar{x}\right) \nabla^{2} P(\bar{x})-P(\bar{x}) \nabla^{2} G\left(\bar{x}_{0}, \bar{x}\right)=P(\bar{x}) \delta\left(\bar{x}-\bar{x}_{0}\right) \tag{7.4.2}
\end{equation*}
$$

This expression is multiplied by a differential element $d \mathcal{V}$ at $\bar{x}$ and integrated over the entire domain. Gradients of products are covered by the same rule as the scalar derivative of a product, and $\nabla^{2} \equiv \nabla \cdot \nabla$, so we have

$$
\begin{align*}
\iiint_{\mathcal{V}} & {\left[G\left(\bar{x}_{0}, \bar{x}\right) \nabla^{2} P(\bar{x})-P(\bar{x}) \nabla^{2} G\left(\bar{x}_{0}, \bar{x}\right)\right] d \mathcal{V} } \\
& \equiv \iiint_{\mathcal{V}} \nabla \cdot\left[G\left(\bar{x}_{0}, \bar{x}\right) \nabla P(\bar{x})-P(\bar{x}) \nabla G\left(\bar{x}_{0}, \bar{x}\right)\right] d \mathcal{V} \\
& =\iiint_{\mathcal{V}} P(\bar{x}) \delta\left(\bar{x}-\bar{x}_{0}\right) d \mathcal{V} \tag{7.4.3}
\end{align*}
$$

The filtering property of a Dirac delta function simplifies the right side. If $\bar{x}_{0}$ is inside $\mathcal{V}$, then the integral will give $P$ at $\bar{x}_{0}$. This is the case that usually interests us, but there also is reason to consider $\bar{x}_{0}$ outside $\mathcal{V}$. Let us denote $\chi\left(\bar{x}_{0}\right)$ as a coefficient that captures both alternatives,

$$
\chi=\left\{\begin{array}{l}
1 \text { if } \bar{x}_{0} \in \mathcal{V}  \tag{7.4.4}\\
0 \text { if } \bar{x}_{0} \notin \mathcal{V}
\end{array}\right.
$$

The left side of Eq. (7.4.3) is described by the divergence theorem. The differential element is situated at $\bar{x}$ within $\mathcal{V}$. To assure that we recognize that the resulting surface integral applies to field points that are on $\mathcal{S}$, we designate those points as $\bar{x}_{s}$ and denote the gradient as $\nabla_{s}$. Also, the usual statement of the divergence theorem uses the normal vector on the surface that is oriented outward from $\mathcal{V}$, but our studies have taken this direction to be oriented into the fluid. Thus, we shall use the divergence theorem with $-\bar{n}\left(\bar{x}_{s}\right)$, which leads to

$$
\begin{equation*}
\iint_{\mathcal{S}}\left(-\bar{n}\left(\bar{x}_{s}\right)\right) \cdot\left[G\left(\bar{x}_{0}, \bar{x}_{s}\right) \nabla_{s} P\left(\bar{x}_{s}\right)-P\left(\bar{x}_{s}\right) \nabla_{s} G\left(\bar{x}_{0}, \bar{x}_{s}\right)\right] d S=\chi\left(\bar{x}_{0}\right) P\left(\bar{x}_{0}\right) \tag{7.4.5}
\end{equation*}
$$

The gradient of $P\left(\bar{x}_{s}\right)$ is related to the complex particle velocity amplitude by Euler's equation, whose application leads to the final form of the Kirchhoff-Helmholtz integral theorem (KHIT),

$$
\begin{align*}
\chi\left(\bar{x}_{0}\right) P\left(\bar{x}_{0}\right)= & \iint_{\mathcal{S}}\left[i \omega \rho_{0} \bar{n}\left(\bar{x}_{s}\right) \cdot \bar{V}\left(\bar{x}_{s}\right) G\left(\bar{x}_{0}, \bar{x}_{s}\right)\right.  \tag{7.4.6}\\
& \left.+P\left(\bar{x}_{s}\right) \bar{n}\left(\bar{x}_{s}\right) \cdot \nabla_{s} G\left(\bar{x}_{0}, \bar{x}_{s}\right)\right] d \mathcal{S}
\end{align*}
$$

This relation has a simple interpretation in terms of point sources. For the first term in the integrand, we observe that $i \omega \rho_{0} \bar{n}\left(\bar{x}_{s}\right) \cdot V\left(\bar{x}_{s}\right) d \mathcal{S}$ is a complex amplitude of the mass acceleration across a differential patch of the boundary. The pressure radiated by a point source having this mass acceleration is $i \omega \rho_{0} \bar{n}\left(\bar{x}_{s}\right) \cdot V\left(\bar{x}_{s}\right) d \mathcal{S} G\left(\bar{x}_{0}, \bar{x}_{s}\right)$. The interpretation of the second term follows from the observation that a similar term occurred in Eq. (6.5.90) for a dipole. A comparison with that term shows that the $-P\left(\bar{x}_{s}\right) \bar{n}\left(\bar{x}_{s}\right) d \mathcal{S} / 4 \pi$ is a differential dipole moment that is oriented perpendicularly
to the surface. Thus, we may regard the KHIT as a statement that a vibrating surface acts as though it was a continuous sheet of monopoles and dipoles.

The KHIT that is found in a text or a research paper might not exactly replicate Eq. (7.4.6). One difference might arise from the use of $\operatorname{Re}(\exp (-i \omega t))$ to represent a harmonic variation. Here, like everywhere else in this book, this alternate convention leads to a complex conjugate representation. Another difference might lie in the definition of the Green's function, which might insert a minus sign and/or omit the $4 \pi$ factor. Differences in the definitions of the Green's function cease to be an issue if the definition is explicitly substituted into the theorem. The free-space Green's function we have defined is

$$
\begin{equation*}
G\left(\bar{x}_{0}, \bar{x}_{s}\right)=\frac{1}{4 \pi \hat{r}} e^{-i k \hat{r}}, \quad \hat{r}=\left|\bar{x}_{s}-\bar{x}_{0}\right| \tag{7.4.7}
\end{equation*}
$$

(The use of $\hat{r}$ to denote the source-field point distance is intended to avoid confusion with $r$ as the radial coordinate.) The meaning of $\nabla_{s} G\left(\bar{x}_{0}, \bar{x}_{S}\right)$ in the integrand is that it is the gradient at $\bar{x}_{s}$ with $\bar{x}_{0}$ held fixed. This gradient was evaluated in Sect. 6.5.1 when we used a Taylor series to derive the dipole field. The direction in which $\hat{r}$ increases most rapidly when $\bar{x}_{0}$ is fixed is the direction of increasing $\hat{r}$. Therefore, the procedure we followed earlier yields

$$
\begin{align*}
\bar{n}\left(\bar{x}_{s}\right) \cdot \nabla_{s} G\left(\bar{x}_{0}, \bar{x}\right) & =\bar{n}\left(\bar{x}_{s}\right) \cdot \nabla_{s} \hat{r} \frac{\partial}{\partial \hat{r}}\left(\frac{1}{4 \pi \hat{r}} e^{-i k \hat{r}}\right) \\
& =-\bar{n}\left(\bar{x}_{s}\right) \cdot\left(\frac{\bar{x}_{s}-\bar{x}_{0}}{\hat{r}}\right)\left(\frac{i k \hat{r}+1}{4 \pi \hat{r}^{2}}\right) e^{-i k \hat{r}} \tag{7.4.8}
\end{align*}
$$

Thus, the KHIT version that explicitly displays the free-space Green's function is

$$
\begin{align*}
\chi\left(\bar{x}_{0}\right) P\left(\bar{x}_{0}\right)=\frac{1}{4 \pi} \iint_{\mathcal{S}} & {\left[i \omega \rho_{0} \bar{n}\left(\bar{x}_{s}\right) \cdot \bar{V}\left(\bar{x}_{s}\right)\right.} \\
& \left.-\bar{n}\left(\bar{x}_{s}\right) \cdot\left(\frac{\bar{x}_{s}-\bar{x}_{0}}{\hat{r}}\right)\left(i k+\frac{1}{\hat{r}}\right) P\left(\bar{x}_{s}\right)\right] \frac{e^{-i k \hat{r}}}{\hat{r}} d \mathcal{S} \tag{7.4.9}
\end{align*}
$$

The KHIT has not been used frequently to study two-dimensional systems. If one wishes to do so, it would be incorrect to employ the preceding form. The proper formulation would use Eq. (7.4.6) with the two-dimensional free-space Green's function in Eq. (7.3.47), which is $G\left(\bar{x}, \bar{x}_{s}\right)=H_{0}(k R) /(4 i)$.

As was noted before we began the derivation, the pressure and particle velocity cannot both be specified on $\mathcal{S}$. Thus, it is reasonable to question whether the KHIT has practical use. Indeed, it does! It may be used to verify a solution that was obtained by another method, but it seldom is. A common application uses the KHIT to evaluate a farfield pressure field when the surface response has been obtained by a numerical simulation technique, or by experimental measurements. An important version of the

KHIT lies in the case not covered by the definition of $\chi\left(\bar{x}_{0}\right)$, Eq. (7.4.4), specifically $\bar{x}_{0}$ being situated on $\mathcal{S}$. In that case, the KHIT gives an equation for the surface pressure at a specific field point on the surface in terms of the surface pressure and velocity distributions. Just as a differential equation relates a variable to its derivatives, this version of the KHIT is an integral equation for the pressure. The solution of this equation by numerical methods is known as a boundary element formulation. This concept is explored in Sect.7.5.2.

EXAMPLE 7.9 Consider a spherical cavity in the case where the container executes a radially symmetric oscillation. In that case, the complex radial velocity on the boundary, $r=a$, is $V_{r}=V_{0}$. This field was analyzed in Sect. 6.3.3. It is desired to use KHIT to verify the analytical solution for the radially symmetric pressure in the interior. To do so, use the analytical solution to determine the pressure at $r=a$, then use that distribution to formulate the KHIT. Evaluate the pressure as a function of $k r$ according to KHIT, and compare it to the analytical solution for the pressure field. The frequency for the evaluation is $k a=\pi$.

## Significance

The formulation of KHIT for a system whose geometry is not too complicated allows us to focus on the meaning of each term. The computed results will suggest an attribute that is crucial to the application of the KHIT as the foundation of a boundary element formulation.

## Solution

Equation (6.3.27) states that the radially symmetric pressure and particle velocity inside a vibrating sphere are

$$
\begin{gather*}
P=-i \rho_{0} c V_{0}\left(\frac{a}{r}\right) \frac{k a \sin (k r)}{k a \cos (k a)-\sin (k a)} \\
V_{r}=V_{0}\left(\frac{a}{r}\right)^{2} \frac{k r \cos (k r)-\sin (k r)}{k a \cos (k a)-\sin (k a)} \tag{1}
\end{gather*}
$$

At $r=a$, these expressions give

$$
\begin{equation*}
P=-i \rho_{0} c V_{0} \frac{k a \sin (k a)}{k a \cos (k a)-\sin (k a)}, \quad V_{r}=V_{0} \tag{2}
\end{equation*}
$$

The task is to use Eq. (2) to form the integrand of the KHIT, then evaluate the pressure within the cavity, and compare it to the pressure in Eq. (1). The field is radially symmetric, so there is no difference in the dependence of $P$ along any radial line. We shall designate this line as the polar axis $z$ because doing so simplifies the formulation. Figure 1 shows a generic surface point $\bar{x}_{s}$ at polar angle $\psi$, and field point $\bar{x}_{0}$ at distance $z_{0}$ on the $z$-axis. (This picture is descriptive of all points at the same value of $\psi$ for any azimuthal angle $\theta$.)


Figure 1.

We could obtain a description of the terms in the KHIT integrand from a trigonometric analysis. However, one objective here is to get an idea of how to formulate the theorem in a general situation where the geometrical configuration is less amenable to a pictorial representation. Thus, all vectorial quantities will be represented in terms of components relative to $x y z$. The normal direction points into the spherical cavity, so the terms appearing in Eqs. (7.4.7) and (7.4.8) are

$$
\begin{gather*}
\bar{x}_{s}=a\left(\sin \psi \bar{e}_{x}+\cos \psi \bar{e}_{z}\right), \quad \bar{x}_{0}=z_{0} \bar{e}_{z} \\
\bar{r}=\left|\bar{x}_{s}-\bar{x}_{0}\right|=\left(a^{2}+z_{0}^{2}-2 a z_{0} \cos \psi\right)^{1 / 2} \\
\bar{n}\left(\bar{x}_{s}\right)=-\bar{e}_{r}=-\sin \psi \bar{e}_{x}-\cos \psi \bar{e}_{z}  \tag{1}\\
-\bar{n}\left(\bar{x}_{s}\right) \cdot \frac{\bar{x}_{s}-\bar{x}_{0}}{\hat{r}}=\cos \beta=\frac{a-z_{0} \cos \psi}{\hat{r}}
\end{gather*}
$$

An axisymmetric differential area element is $2 \pi(a \sin \psi)(a d \psi)$, so we have characterized the geometric variables required to form the KHIT. For $P\left(\bar{x}_{s}\right)$, we use the first of Eq. (2). Because the normal direction is defined to point into the fluid domain, the surface velocity is $\bar{n}\left(\bar{x}_{s}\right) \cdot \bar{V}\left(\bar{x}_{s}\right)=-V_{r}=-V_{0}$. All parameters are specified in nondimensional form, so we shall replace $a, \hat{r}$, and $z_{0}$ with $k a, k \hat{r}$, and $k z_{0}$ everywhere.

The descriptions are substituted into the KHIT, with the result that we must evaluate

$$
\begin{equation*}
\frac{\chi P(\bar{x})}{\rho_{0} c v_{0}}=-\frac{i}{2}(k a)^{2}\left[\mathcal{I}_{1}+\frac{k a \sin (k a)}{k a \cos (k a)-\sin (k a)} \mathcal{I}_{2}\right] \tag{2}
\end{equation*}
$$

where

$$
\begin{gather*}
\mathcal{I}_{1}=\int_{0}^{\pi} \frac{e^{-i k \hat{r}}}{k \hat{r}} \sin \psi d \psi \\
\mathcal{I}_{2}=\int_{0}^{\pi} \frac{e^{-i k \hat{r}}}{k \hat{r}}(\cos \beta)(\sin \psi) d \psi \tag{3}
\end{gather*}
$$

Equation (1) give $k \hat{r}$ and $\cos \beta$ as functions of $\psi$ for specified values of $k a$ and $k z_{0}$, so we proceed to the evaluation. If we were to search through tables of integrals, we might find expressions for $\mathcal{I}_{1}$ and $\mathcal{I}_{2}$. However, the idea here is to use numerical methods because we would have no expectation of an analytical result for more complicated shapes.

The data in Fig. 2 was obtained with MATLAB. For each value of $k z_{0}$, the anonymous function capability was used to define kr _hat as a function of psi, which then was used to define cos_beta as a function of psi. Both were then used to define the integrands of $\mathcal{I}_{1}$ and $\mathcal{I}_{2}$, and the integrals were found by invoking the quadl routine. The error tolerance was decreased to $10^{-8}$.


Figure 2.

According to Eq. (2), $P /\left(\rho_{0} c V_{a}\right)$ within the cavity is purely imaginary. The value obtained from the KHIT nowhere differs from the analytical value by more than $1.4\left(10^{-7}\right)$, and the real part is no greater than $\pm 1.8\left(10^{-10}\right)$. The discontinuity at $k z_{0}=4$ results because the pressure obtained from KHIT outside the domain should be zero. For $k z_{0}>k a$, the computed value is less than $6\left(10^{-8}\right)$ out to $k z_{0}=1.5 \mathrm{ka}$.

The evaluation skipped $k z_{0}=k a$ because both integrands in that case are singular at $\psi=0$. Nevertheless, it is interesting to wonder what would happen if $k z_{0}$ were very close to $k a$. To test this, a sequence of evaluations with $k z_{0}=k a \mp 10^{-n}$ were carried out. The results are tabulated below.

| $p /\left(\rho_{0} c v_{0}\right)$ <br> $k z_{0}-k a$ | Analytical |  |
| :---: | :---: | :---: |
| $-\left(10^{-5}\right)$ | $-5.1281\left(10^{-9}\right)-1.6295 i$ | $-1.6295 i$ |
| $-\left(10^{-6}\right)$ | $-5.1282\left(10^{-9}\right)-0.8146 i$ | $-1.6295 i$ |
| $-\left(10^{-7}\right)$ | $-5.1282\left(10^{-9}\right)-0.8145 i$ | $-1.6295 i$ |
| $-\left(10^{-8}\right)$ | Singular | $-1.6295 i$ |
| $\left(10^{-8}\right)$ | Singular | 0 |
| $\left(10^{-7}\right)$ | $-5.1282\left(10^{-9}\right)-0.8144 i$ | 0 |
| $\left(10^{-6}\right)$ | $-5.1282\left(10^{-9}\right)-0.8143 i$ | 0 |
| $\left(10^{-5}\right)$ | $-5.1283\left(10^{-9}\right)-0.8127 i$ | 0 |
| $\left(10^{-4}\right)$ | $-5.1288\left(10^{-9}\right)+6.5779\left(10^{-7}\right) i$ | 0 |

The table indicates that as the field point approaches the surface from the inside, the pressure drops from the analytical value to approximately half that value, but it is singular at the closest distance. The trend for points outside the sphere begins with a singularity at the closest distance, then approximately half the value on the surface, then zero. This behavior will be examined in our development of boundary elements, where it will be proven that $\chi\left(\bar{x}_{0}\right)=0.5$ if $\bar{x}_{0}$ actually is situated on a smooth surface.

### 7.4.2 Acoustic Radiation into an Exterior Domain

This situation we consider here is one where a vibrating surface $\mathcal{S}$ radiates into a surrounding fluid domain $\mathcal{V}$, and the extent of $\mathcal{V}$ is infinite. We cannot apply KHIT directly because its derivation required that the vibrating boundary encloses the domain. However, the derivation of KHIT did not require that the boundary be continuous. Consequently, it is valid if we consider the boundary of the fluid domain to be the composite of the radiating surface $\mathcal{S}$ and a very large sphere $\mathcal{S}_{O}$ having radius $r_{O}$ that encloses $\mathcal{S}$. The fluid domain $\mathcal{V}$ is the region between $\mathcal{S}$ and $\mathcal{S}_{O}$. Figure 7.18 depicts the decomposition of the region. After we modify KHIT to describe this configuration, the limit as $r_{O} \rightarrow \infty$ will provide the theorem we seek.

Fig. 7.18 Acoustic region
for derivation of the Kirchhoff-Helmholtz integral theorem describing radiation from vibrating body $\mathcal{S}$. The surrounding sphere is virtual


The integral in the KHIT now becomes the sum of integrals over the vibrating surface $\mathcal{S}$ and the virtual $\mathcal{S}_{O}$. The first has the same form as the left side of Eq.(7.4.6). The only difference stems from the convention that $\bar{n}\left(\bar{x}_{s}\right)$ always is oriented into the fluid. Hence, for the radiation problem, it is outward from $\mathcal{S}$, whereas it is oriented inward when $\mathcal{S}$ is the outer boundary of a cavity. The other part of the surface integral is the contribution from $\mathcal{S}_{O}$. The normal direction for this portion of the boundary is $-\bar{e}_{r}$, so that $\bar{n}\left(\bar{x}_{s}\right) \cdot \nabla=-\partial / \partial r$ and $\bar{n}\left(\bar{x}_{s}\right) \cdot \bar{V}\left(\bar{x}_{s}\right)=-V_{r}$. Also, in terms of spherical coordinates, we have $d \mathcal{S}_{O}=r_{O}^{2} \sin \psi d \theta d \psi$. Thus, the contribution of the surrounding sphere is

$$
\begin{align*}
\iint_{\mathcal{S}_{O}} & {\left[i \omega \rho_{0} \bar{n}\left(\bar{x}_{s}\right) \cdot V\left(\bar{x}_{s}\right) G\left(\bar{x}_{0}, \bar{x}_{s}\right)+P\left(\bar{x}_{s}\right) \bar{n}\left(\bar{x}_{s}\right) \cdot \nabla_{s} G\left(\bar{x}_{0}, \bar{x}_{s}\right)\right] d S } \\
& =\int_{0}^{\pi} \int_{-\pi}^{\pi}\left[-i \omega \rho_{0} V_{r}\left(\bar{x}_{s}\right) G\left(\bar{x}_{0}, \bar{x}_{s}\right)-P\left(\bar{x}_{s}\right) \frac{\partial}{\partial r} G\left(\bar{x}_{0}, \bar{x}_{s}\right)\right] r_{O}^{2} \sin \psi d \theta d \psi \tag{7.4.10}
\end{align*}
$$

Our study of radiated power in Sect.4.5.3, which led to the Sommerfeld radiation condition for a three-dimensional domain, proved that if $r_{O}$ is very large, then the pressure should exhibit spherical spreading. The specific description was

$$
\begin{equation*}
P=\frac{1}{r_{O}} f(\psi, \theta) e^{-i k r_{O}}+O\left(\frac{1}{r_{O}^{2}}\right) \tag{7.4.11}
\end{equation*}
$$

Application of Euler's equation shows that the radial velocity is

$$
\begin{equation*}
V_{r}=\frac{1}{\rho_{0} c r_{O}} f(\psi, \theta) e^{-i k r_{O}}+O\left(\frac{1}{r_{O}^{2}}\right) \tag{7.4.12}
\end{equation*}
$$

where the two variables have been written separately, rather than as $P=\rho_{0} c V_{r}$ to emphasize that this property is only satisfied to the leading order in $r_{O}$. The Green's function $G\left(\bar{x}_{s}, \bar{x}_{0}\right)$ also will satisfy the farfield approximation when $\bar{x}_{s}$ is situated on $S_{O}$. This is so because the source location $\bar{x}_{0}$ is at a fixed finite distance from the center of the sphere, and $r_{O}$ can be as large as necessary to make this statement be true. Thus, it must be that for any $\bar{x}_{s} \in \mathcal{S}_{O}$,

$$
\begin{align*}
G\left(\bar{x}_{0}, \bar{x}_{s}\right) & =\frac{1}{r_{O}} g(\psi, \theta) e^{-i k r_{O}}+O\left(\frac{1}{r_{O}^{2}}\right) \\
\frac{\partial}{\partial r} G\left(\bar{x}_{0}, \bar{x}_{s}\right) & =-\frac{i k}{r_{O}} g(\psi, \theta) e^{-i k r_{O}}+O\left(\frac{1}{r_{O}^{2}}\right) \tag{7.4.13}
\end{align*}
$$

When these farfield representations are substituted into Eq.(7.4.10), the terms that are products of the $f$ and $g$ functions cancel. What remains in the integrand is $O\left(1 / r_{O}^{3}\right)$. This term is multiplied by $r_{O}^{2}$ in the differential area $d \mathcal{S}$, so the quantity that is integrated is $O\left(1 / r_{O}\right)$. Consequently, the integral vanishes in the limit as $r_{O} \rightarrow \infty$. Therefore, the KHIT only requires integration over $\mathcal{S}$.

The consequence is that Eq. (7.4.6) also applies when $\mathcal{V}$ is the infinite fluid regions surrounding a vibrating surface $S$. It is helpful to restate the theorem. The general form leaves the Green's function unspecified, so it is valid for two- and threedimensional problems. This description is

$$
\begin{align*}
\chi\left(\bar{x}_{0}\right) P\left(\bar{x}_{0}\right)= & \iint_{\mathcal{S}}\left[i \omega \rho_{0} \bar{n}\left(\bar{x}_{s}\right) \cdot \bar{V}\left(\bar{x}_{s}\right) G\left(\bar{x}_{0}, \bar{x}_{s}\right)\right.  \tag{7.4.14}\\
& \left.+P\left(\bar{x}_{s}\right) \bar{n}\left(\bar{x}_{s}\right) \cdot \nabla_{s} G\left(\bar{x}_{0}, \bar{x}_{s}\right)\right] d \mathcal{S}
\end{align*}
$$

The usual application is to three-dimensional systems. Explicitly, representing the free-space Green's function in the preceding leads to

$$
\begin{align*}
\chi\left(\bar{x}_{0}\right) P\left(\bar{x}_{0}\right)= & -\frac{1}{4 \pi} \iint_{\mathcal{S}}\left[\bar{n}\left(\bar{x}_{s}\right) \cdot\left(\frac{\bar{x}_{s}-\bar{x}_{0}}{\hat{r}}\right)\left(i k+\frac{1}{\hat{r}}\right) P\left(\bar{x}_{s}\right)\right.  \tag{7.4.15}\\
& \left.-i \omega \rho_{0} \bar{n}\left(\bar{x}_{s}\right) \cdot \bar{V}\left(\bar{x}_{s}\right)\right] \frac{e^{-i k \hat{r}}}{\hat{r}} d \mathcal{S}, \quad \hat{r}=\left|\bar{x}_{s}-\bar{x}_{0}\right|
\end{align*}
$$

This is the form in which KHIT typically appears.
The fact that KHIT has the same appearance for the interior and exterior problem can cause confusion. Compare the following situations: (1) Surface $\mathcal{S}$ is the exterior boundary of a cavity, (2) the same surface $\mathcal{S}$ is the vibrating surface of a transducer surrounded by fluid. The normal velocity distribution in both systems is the same. So what is the difference in KHIT for these systems? The answer lies in the definition of the normal direction on the surface, $\bar{n}\left(\bar{x}_{s}\right)$ as the normal that points into the fluid domain. Thus, if $\mathcal{S}$ is the surface that encloses a cavity, $n(\bar{x})$ is oriented onto the region bounded by $\mathcal{S}$. In contrast, if $S$ is the surface of a body that is surrounded by fluid, then $n(\bar{x})$ points outward from the region bounded by $\mathcal{S}$. In other words, $\bar{n}\left(\bar{x}_{s}\right)$ is reversed between the two cases. The only other change is that $\chi\left(\bar{x}_{0}\right)$ is one if $\bar{x}_{0}$ is situated within the cavity, or if it is exterior to the transducer, and $\chi\left(\bar{x}_{0}\right)$ is zero if $\bar{x}_{0}$ is outside the cavity or inside the transducer.

As was the case for the interior cavity, the KHIT is not a solution to a problem because we cannot specify both the pressure and normal velocity at a location. A common use is to determine the pressure, especially in the farfield, when the surface response has been determined by another method. Example 7.10, which follows, describes such an application in conjunction with an approximation of the surface pressure.

If only the farfield pressure is of interest, the KHIT may be written in a form that is somewhat simpler. This is done by introducing the same approximations as those employed to describe the farfield of a collection of point sources. The origin is inside $\mathcal{S}$. As part of the restriction to the farfield, the distance to $\bar{x}_{0}$ from the origin must be much greater than the maximum distance from the origin to a point on the surface, that is, $\left|\bar{x}_{0}\right| \gg \max \left(\left|\bar{x}_{s}\right|\right)$ In that case, the line from $\bar{x}_{S}$ to $\bar{x}_{0}$, which is $\bar{x}_{0}-\bar{x}_{s}$, will approach parallelism to $\bar{x}_{0}$. Thus, the distance $\hat{r}$ from the surface point to the field point may be approximated as the radial distance $r=\left|\bar{x}_{0}\right|$ from the origin less the projection of $\bar{x}_{s}$ onto this radial line. The unit vector for the radial line is $\bar{x}_{0} / r$, so we have

$$
\begin{equation*}
r=\left|\bar{x}_{0}\right|, \quad \hat{r} \equiv\left|\bar{x}_{s}-\bar{x}_{0}\right| \approx r-\bar{x}_{s} \cdot \frac{\bar{x}_{0}}{r} \tag{7.4.16}
\end{equation*}
$$

where the approximation becomes increasingly accurate with increasing $\left|\bar{x}_{0}\right|$. We use this approximation to factorize the phase of the free-space Green's function and its normal derivative in Eq. (7.4.8). The result is that

$$
\begin{align*}
& G\left(\bar{x}_{s}, \bar{x}_{0}\right) \approx\left(\frac{e^{-i k r}}{4 \pi r}\right) e^{i k \bar{x}_{s} \cdot \bar{x}_{0} / r} \\
& \bar{n}\left(\bar{x}_{s}\right) \cdot \nabla_{s} G\left(\bar{x}_{s}, \bar{x}\right) \approx \bar{n}\left(\bar{x}_{s}\right) \cdot \frac{\bar{x}_{0}}{r}\left(\frac{i k e^{-i k r}}{4 \pi r}\right) e^{i k \bar{x}_{s} \cdot \bar{x}_{0} / r} \tag{7.4.17}
\end{align*}
$$

The corresponding form of the KHIT is

$$
\begin{equation*}
P_{\mathrm{ff}}\left(\bar{x}_{0}\right)=i k \frac{e^{-i k r}}{4 \pi r} \iint_{\mathcal{S}}\left[\rho_{0} c \bar{V}\left(\bar{x}_{s}\right) \cdot \bar{n}\left(\bar{x}_{s}\right)+P\left(\bar{x}_{s}\right) \bar{n}\left(\bar{x}_{s}\right) \cdot \frac{\bar{x}_{0}}{r}\right] e^{i k \bar{x}_{s} \cdot \bar{x}_{0} / r} d S \tag{7.4.18}
\end{equation*}
$$

The sole dependence of the integrand on the location of the field point is the value of $r$ and the unit vector $\bar{x}_{0} / r$. When we use the polar angle $\psi_{0}$ and azimuthal angle $\theta_{0}$ to locate that point, the farfield version of KHIT reduces to

$$
\begin{equation*}
P_{\mathrm{ff}}\left(\bar{x}_{0}\right)=i k \frac{e^{-k r}}{4 \pi r} f\left(\psi_{0}, \theta_{0}\right) \tag{7.4.19}
\end{equation*}
$$

This is the same form as that in Eq. (4.5.43). The difference is that the KHIT tells us how to evaluate the directivity factor, $f\left(\psi_{0}, \theta_{0}\right)$. Of course, this assumes that we know the pressure and normal velocity on the radiating body's surface.

As was true for the interior problem, $G\left(\bar{x}_{0}, \bar{x}_{s}\right)$ need not be the free-space Green's function, but it usually is because any other is difficult to obtain. However, one notable exception is the Green's function used to describe radiation from an infinite planar boundary. Chapter 8 is devoted to this topic. The foundation for that study is the Rayleigh integral, which essentially is the KHIT with a Green's function that is derived from application of the method of images.

EXAMPLE 7.10 A cylindrical tank of length $L$ is submerged. Pulsations in the feed line result in internal pressure fluctuations that induce an axisymmetric vibration of the cylinder. The velocity of the cylinder's surface in the normal direction is $v_{R}=\operatorname{Re}\left(v_{0} \cos (\pi z / L) e^{i \omega t}\right)$ for $-L / 2 \leq z \leq L / 2$, which is a single lobe of a sine function that extends over the length of the cylinder. The ends of the cylinder are immobile. It is desired to determine the farfield pressure distribution. The radiation impedance of an infinite cylinder provides an approximation of the surface pressure generated by the specified velocity. Use this approximation to formulate the farfield pressure according to the KHIT. Evaluate the result for the directivity $(r / a)\left|P_{\mathrm{ff}}\left(\bar{x}_{0}\right)\right| /\left(\rho_{0} c v_{0}\right)$ as function of the spherical angle locating field point $\bar{x}_{0}$. Parameters for this evaluation are $k a=1$ and 6 , and $L / a=4$ and 20.

## Significance

In addition to showing how the KHIT may be formulated for three-dimensional bodies, this example has the purpose of showing that in some occasions, one may use knowledge of fundamentals to construct approximate solutions to problems that otherwise would require sophisticated simulation software.

## Solution

The analysis is based on the suggestion in Example 7.7 that the field close to a cylinder might be reasonably well predicted by a model in which the cylinder is infinite. It is an approximation to ignore the fact that the helical surface waves do not propagate to infinity. In addition, deformation of the cylinder would result in movement of the ends, even if they are made very rigid. Furthermore, even if the ends did not move, the pressure on the ends would not be zero. Nevertheless, the analysis we will pursue is useful because it enables us to understand some phenomena and it provides an order of magnitude check for more exact analyses. Some might refer to this as a "back of the envelope" analysis, but that description is only appropriate in comparison with the effort required for a faithful simulation.

The decomposition of the given radial velocity into its constituent helical surface waves is

$$
\begin{equation*}
v_{R}=\operatorname{Re}\left[\frac{v_{0}}{2}\left(e^{i \mu z}+e^{-i \mu z}\right) e^{i \omega t}\right], \quad \mu=\frac{\pi}{L} \tag{1}
\end{equation*}
$$

Both are axisymmetric, $n=0$. In comparison with Eq.(7.3.16), this motion is a superposition of two helical waves whose velocity coefficients are $V_{1,0}=V_{-1,0}=$ $v_{0} / 2$. The radial wavenumber for both is

$$
\begin{equation*}
\kappa a=\left((k a)^{2}-\pi^{2} \frac{a^{2}}{L^{2}}\right)^{1 / 2} \tag{2}
\end{equation*}
$$

All combinations of the specified values of $k a$ and $L / a$ lead to $\kappa a$ being real, so the helical waves are supersonic. The radiation impedance in Eq.(7.3.35) does not depend on the sign of the axial wavenumber, so it is the same for both helical waves. Adding the contribution of each leads to the complex amplitude of surface pressure being

$$
\begin{equation*}
P=\rho_{0} c Z_{\mathrm{rad}} v_{0} \cos (\mu z), \quad Z_{\mathrm{rad}}=-i \frac{k a}{\kappa a} \frac{H_{0}(\kappa a)}{H_{0}^{\prime}(\kappa a)} \tag{3}
\end{equation*}
$$

The next step is to describe the geometric variables in the farfield representation of the KHIT, Eq. (7.4.18). We do so by using a set of $x y z$ axes to describe vector components. The variables to be represented appear in Fig. 1.


Figure 1.

The field point $\bar{x}_{0}$ has been placed on the $x z$ plane at polar angle $\psi_{0}$. This placement will yield a general result because an axisymmetric surface vibration generates a field that also is axisymmetric. The surface point has cylindrical coordinates $(a, \theta, z)$, and the transverse direction $\bar{e}_{R}$ is normal to the surface. It is stated that the ends of the cylinder are immobile. We shall assume that the pressure on the ends also is zero. (This assumption often is based on the conjecture that radiation from the ends is unimportant if $L / a$ is not small.) The quantities of interest are

$$
\begin{gathered}
\bar{x}_{0}=r \sin \psi_{0} \bar{e}_{x}+r \cos \psi_{0} \bar{e}_{z} \\
\bar{x}_{s}=a \cos \theta \bar{e}_{x}+a \sin \theta \bar{e}_{y}+z z \bar{e}_{z} \\
\bar{n}\left(\bar{x}_{s}\right)=\bar{e}_{R}=\cos \theta \bar{e}_{x}+\sin \theta \bar{e}_{y}
\end{gathered}
$$

These variables depend on $\theta$, so we use $(a d \theta) d z$ as the differential area element and integrate over $-\pi<\theta \leq \pi,-L / 2 \leq z \leq L / 2$. Substitution of the preceding terms and $P=Z_{\mathrm{rad}} V_{R}$ into Eq.(7.4.18) leads to

$$
\begin{align*}
\frac{P_{\mathrm{ff}}\left(\bar{x}_{0}\right)}{\rho_{0} c v_{0}}= & \frac{i k a}{4 \pi r} e^{-i k r} \int_{-L / 2}^{L / 2} \int_{-\pi}^{\pi} \cos (\mu z)(1  \tag{4}\\
& \left.+Z_{\mathrm{rad}} \cos \theta \sin \psi_{0}\right) e^{i k\left(a \cos \theta \sin \psi_{0}+z \cos \psi_{0}\right)} d \theta d z
\end{align*}
$$

The exponential may be split into factors that depend solely on $\theta$ and $z$, and $Z_{\mathrm{rad}}$ is independent of $\theta$ and $z$. Furthermore, the $\theta$ integral depends only on $\cos \theta$, so it is the same as twice the integral over $0<\theta \leq \pi$. Recognition of these features reduces Eq. (4) to

$$
\begin{equation*}
\frac{P_{\mathrm{ff}}\left(\bar{x}_{0}\right)}{\rho_{0} c v_{0}}=\frac{i k a}{4 \pi r} e^{-i k r}\left(\mathcal{I}_{1}+\mathcal{I}_{2} Z_{\mathrm{rad}} \sin \psi_{0}\right) \mathcal{I}_{z} \tag{5}
\end{equation*}
$$

Integration over $z$ is described by $\mathcal{I}_{z}$. It is

$$
\begin{aligned}
\mathcal{I}_{z} & =\int_{-L / 2}^{L / 2} \cos (\mu z) e^{i k z \cos \psi_{0}} d z \\
& =\frac{\sin \left[\left(k L \cos \psi_{0}+\mu L\right) / 2\right]}{\left(k \cos \psi_{0}+\mu\right)}+\frac{\sin \left[\left(k L \cos \psi_{0}-\mu L\right) / 2\right]}{\left(k \cos \psi_{0}-\mu\right)}
\end{aligned}
$$

The $\theta$ integration is described by $\mathcal{I}_{1}$ and $\mathcal{I}_{2}$. They appear in Eq. (9.1.21) in Abramowitz and Stegun as

$$
\begin{gathered}
\int_{0}^{\pi} e^{i k a \cos \theta \sin \psi_{o}} d \theta=\pi J_{0}\left(k a \sin \psi_{0}\right) \\
\int_{0}^{\pi} e^{i k a \cos \theta \sin \psi_{o}} \cos \theta d \theta=i \pi J_{1}\left(k a \sin \psi_{0}\right)
\end{gathered}
$$

Substitution of these integrals into the Eq. (4) leads to

$$
\begin{align*}
\frac{P_{\mathrm{ff}}\left(\bar{x}_{0}\right)}{\rho_{0} c v_{0}}= & \frac{i k L}{2}\left(\frac{a}{r}\right) e^{-i k r}\left[J_{0}\left(k a \sin \psi_{0}\right)+i Z_{\mathrm{rad}} \sin \psi_{0} J_{1}\left(k a \sin \psi_{0}\right)\right]  \tag{6}\\
& \times\left\{\frac{\sin \left[\left(k L \cos \psi_{0}+\mu L\right) / 2\right]}{\left(k L \cos \psi_{0}+\mu L\right)}+\frac{\sin \left[\left(k L \cos \psi_{0}-\mu L\right) / 2\right]}{\left(k L \cos \psi_{0}-\mu L\right)}\right\}
\end{align*}
$$

It should be noted that the farfield pressure decays reciprocally to the radial distance. This must be so because no matter how long the cylinder is, the farfield specification places the field point sufficiently far away that the cylinder seems to occupy a single point.

For each combination of $k a$ and $L / a$, we compute $(r / a)\left|P_{\mathrm{ff}}\right| /\left(\rho_{0} c v_{0}\right)$ as a function of $\psi_{0}$. (The parameter $k L$ is written as $k a(L / a)$ for this computation.) Figure 2 displays polar plots for the four parameter combinations. Proper interpretation of these plots requires that one recognize that because of the axisymmetry of the system, these plots describe the view in any plane that contains the cylinder's axis. The overall picture is that only the shorter cylinder at the lower frequency radiates strongly in all directions. This is not surprising because our earlier studies have indicated that at very low frequencies, bodies radiate as a monopole whose strength depends on the complex amplitude of the volume velocity. For the longer cylinder in both cases, as well as the shorter cylinder at high frequency, the radiation is confined to a small range of angles around $\psi_{0}=90^{\circ}$, which is said to be "broadside" or "beam aspect," both of which are nautical in their origin. Another trend displayed by the graphs is that increasing the frequency increases the maximum farfield pressure.


Figure 2.

A fault of polar plots is that they obscure data close to the origin. For this reason, the same data is plotted in Fig. 3 in abscissa-ordinate form. These graphs vividly show how narrowly the radiation is aimed in all cases except the short cylinder at the low frequency. In the intermediate cases, there are many side lobes, but they are much smaller than the main lobe. The side lobes for the long cylinder at the high frequency are negligibly small.


Figure 3.

## Multipole Expansion of the KHIT

The initial discussion noted that the KHIT essentially describes the pressure field as the superposition of monopoles and dipoles on the vibrating surface. However, the description of the directivity derived in the previous section shows no evidence of a set of sources. The first effort in the following will resolve this apparent conflict. By itself, the development will merely provide a different perspective. However, further analysis based on a restriction to acoustically compact radiators (size substantially less than a wavelength) will lead to a recipe for actually evaluating the pressure field.

General Description We begin with a Taylor series expansion of the Green's function terms. A central point $C$ inside the surface $\mathcal{S}$, such as the centroid of the vibrating body, is defined to be a reference location. Then, the location of any point on $\mathcal{S}$ is described by the relative position $\bar{\xi}$, which is

$$
\begin{equation*}
\bar{\xi}=\bar{x}_{s}-\bar{x}_{C} \tag{7.4.20}
\end{equation*}
$$

Because $\bar{x}_{C}$ is a specified fixed point, we may consider the surface response to be functions of $\bar{\xi}$. An explicit description of the Green's function as a function of $\bar{\xi}$ results from application of Eq. (6.5.82). Doing so gives

$$
\begin{equation*}
G\left(\bar{x}_{0}, \bar{x}_{s}\right)=G\left(\bar{x}_{0}, \bar{x}_{C}\right)+\bar{\xi} \cdot \nabla_{C} G\left(\bar{x}_{0}, \bar{x}_{C}\right)+\frac{1}{2}\left(\bar{\xi} \cdot \nabla_{C}\right)^{2} G\left(\bar{x}_{0}, \bar{x}_{C}\right)+\cdots \tag{7.4.21}
\end{equation*}
$$

where the notation $\nabla_{C}$ serves to emphasize that the gradient is evaluated at $\bar{x}_{C}$. The normal derivative of $G\left(\bar{x}_{0}, \bar{x}_{s}\right)$ in the KHIT is taken at $\bar{x}_{s}$. To put it on the same basis as the above representation of $G\left(\bar{x}_{0}, \bar{x}_{s}\right)$, we invoke Eq.(6.5.86), which states that $\nabla_{s} G\left(\bar{x}_{0}, \bar{x}_{s}\right)=-\nabla_{0} G\left(\bar{x}_{0}, \bar{x}_{C}\right)$. For the present purpose, it is adequate to truncate Taylor series at second derivatives, so we use the preceding to represent the Green's function. Doing so yields

$$
\begin{align*}
\bar{n}(\bar{\xi}) \cdot \nabla_{s} G\left(\bar{x}_{0}, \bar{x}_{s}\right)= & {\left[-\bar{n}(\bar{\xi}) \cdot \nabla_{0}\right]\left[G\left(\bar{x}_{0}, \bar{x}_{C}\right)+\bar{\xi} \cdot \nabla_{C} G\left(\bar{x}_{0}, \bar{x}_{C}\right)+\cdots\right] } \\
= & -\bar{n}(\bar{\xi}) \cdot \nabla_{0} G\left(\bar{x}_{0}, \bar{x}_{C}\right) \\
& +\left(\bar{n}(\bar{\xi}) \cdot \nabla_{0}\right)\left(\bar{\xi} \cdot \nabla_{0}\right) G\left(\bar{x}_{0}, \bar{x}_{C}\right)+\cdots \tag{7.4.22}
\end{align*}
$$

The dipole and quadrupole strengths were described in Chap. 6 in terms of derivatives at the field point, with the source point defined to be the origin of a Cartesian coordinate system. Therefore, we convert the preceding expression to scalar form by explicitly describing the gradients. In the following, $x_{1}, x_{2}, x_{3}$ are the Cartesian coordinates of $\bar{x}_{0}$. (In most situations, we use the free-space Green's function and designate $\bar{x}_{C}$ as the origin.) Substitution of the series expansion of $G\left(\bar{x}_{0}, \bar{x}_{s}\right)$ and its gradient into the KHIT, followed by collection of the coefficients of each order of derivative, yields

$$
\begin{align*}
P\left(\bar{x}_{0}\right)= & 4 \pi\left[A G\left(\bar{x}_{0}, \bar{x}_{C}\right)+\sum_{j=1}^{3} D_{j} \frac{\partial}{\partial x_{j}} G\left(\bar{x}_{0}, \bar{x}_{C}\right)\right.  \tag{7.4.23}\\
& \left.+\sum_{j=1}^{3} \sum_{m=1}^{3} Q_{j, m} \frac{\partial^{2}}{\partial x_{j} \partial x_{m}} G\left(\bar{x}_{0}, \bar{x}_{C}\right)+\cdots\right]
\end{align*}
$$

The derivatives of the Green's function are described by Eq. (6.5.123). The source strength, dipole moments, and quadrupole strengths appearing in this expression are

$$
\begin{gather*}
A=\frac{1}{4 \pi} \iint_{\mathcal{S}} i k \rho_{0} c \bar{V}(\bar{\xi}) \cdot \bar{n}(\bar{\xi}) d \mathcal{S} \\
D_{j}=-\frac{1}{4 \pi} \iint_{\mathcal{S}}\left[i k \rho_{0} c \bar{V}(\bar{\xi}) \cdot \bar{n}(\bar{\xi}) \xi_{j}+P(\bar{\xi}) n_{j}(\bar{\xi})\right] d \mathcal{S} \\
\hat{Q}_{j, m}=\frac{1}{8 \pi} \iint_{\mathcal{S}}\left\{i k \rho_{0} c \bar{V}(\bar{\xi}) \cdot \bar{n}(\bar{\xi}) \xi_{j} \xi_{m}+P(\bar{\xi})\left[n_{j}(\bar{\xi}) \xi_{m}+n_{m}(\bar{\xi}) \xi_{j}\right]\right\} d \mathcal{S} \tag{7.4.24}
\end{gather*}
$$

The expression for $\hat{Q}_{j, m}$ has a symmetric form because a mixed derivative does not depend on the order of differentiation. The total strength of the lateral quadrupoles is $Q_{j, m} \equiv 2 \hat{Q}_{m, j}$.

By itself, this representation generally is not useful because convergence will require many higher order poles. The exception is low frequencies, in which case, the body is compact. That is, its largest dimension $a$ is much smaller than the acoustic wavelength, so $k a \ll 1$. In such situations, the preceding should be adequate. Indeed, the monopole contribution, whose strength is proportional to the volume velocity of the source, often is sufficient.

A Vibrating Rigid Body The multipole expansion provides an interesting perspective when the object from which the signal emanates is a rigid body. We restrict our attention to acoustically compact bodies, or equivalently, to very low frequencies. The theorems of kinematics state that the movement of such an object is the combination of a translation that follows an arbitrary point in the body, and a rotation in which the arbitrary point does not move. We shall use the centroid $C$ of the body as the arbitrary point, and define this point to be the origin of the coordinate system.

The translational velocity $\bar{v}_{C}$ and angular velocity $\bar{\Omega}$ are taken to be harmonic at the same frequency. The complex velocity amplitude of any point in the body is

$$
\begin{equation*}
\bar{V}=\bar{V}_{C}+\bar{\Omega} \times \bar{x} \tag{7.4.25}
\end{equation*}
$$

The normal velocity on the surface is found by evaluating this expression at surface point $\bar{x}_{s}$, then taking the dot product with the surface normal $\bar{n}\left(\bar{x}_{s}\right)$.

The result is used to evaluate the source strengths in Eq. (7.4.24). For the monopole amplitude, we have

$$
\begin{equation*}
A=\frac{1}{4 \pi} i k \rho_{0} c \iint_{\mathcal{S}} \bar{n}(\bar{\xi}) \cdot\left[\bar{V}_{C}+(\bar{\Omega} \times \bar{\xi})\right] d \mathcal{S} \tag{7.4.26}
\end{equation*}
$$

We apply the divergence theorem to this expression, and recognize that $\bar{V}_{C}$ is independent of position, so that $\nabla \cdot \bar{V}_{C}=0$. In addition, $\nabla \cdot(\bar{\Omega} \times \bar{x})$ is identically zero. Consequently, $A=0$. The monopole amplitude is zero because the nature of rigid body motion is that outward movement of some regions of the surface is balanced by inward motion of other regions, so the net volume velocity is zero.

The dipole moment obtained from Eq. (7.4.24) is

$$
\begin{align*}
D_{j}= & -\frac{1}{4 \pi} i k \rho_{0} c \iint_{\mathcal{S}}\left[\bar{V}_{C} \xi_{j} \cdot \bar{n}(\bar{\xi})+(\bar{\Omega} \times \bar{\xi}) \xi_{j} \cdot \bar{n}(\bar{\xi})\right] d \mathcal{S} \\
& -\frac{1}{4 \pi} \iint_{\mathcal{S}} P(\bar{\xi}) n_{j}(\bar{\xi}) d \mathcal{S}, \quad j=1,2,3 \tag{7.4.27}
\end{align*}
$$

We use the divergence theorem to convert the first integral to a volume integral,

$$
\begin{align*}
D_{j}= & -\frac{1}{4 \pi} i k \rho_{0} c \iiint_{\mathcal{V}}\left[\nabla \cdot\left(\bar{V}_{C} \xi_{j}\right)+\nabla \cdot\left((\bar{\Omega} \times \bar{\xi}) \xi_{j}\right)\right] d \mathcal{V} \\
& -\frac{1}{4 \pi} \iint_{\mathcal{S}} P(\bar{\xi}) n_{j}(\bar{\xi}) d \mathcal{S}, \quad j=1,2,3 \tag{7.4.28}
\end{align*}
$$

Both terms in the first integrand have the form $\nabla \cdot\left(\bar{U} \xi_{j}\right)$. The gradient is taken at the surface point. Thus, $\nabla=\bar{e}_{1}\left(\partial / \partial \xi_{1}\right)+\bar{e}_{2}\left(\partial / \partial \xi_{2}\right)+\bar{e}_{3}\left(\partial / \partial \xi_{3}\right)$. Expanding $\bar{U}$ into its components gives

$$
\begin{equation*}
\nabla \cdot\left(\bar{U} \xi_{j}\right)=\frac{\partial}{\partial \xi_{1}}\left(U_{1} \xi_{j}\right)+\frac{\partial}{\partial \xi_{2}}\left(U_{2} \xi_{j}\right)+\frac{\partial}{\partial \xi_{3}}\left(U_{3} \xi_{j}\right)=U_{j}+\xi_{j} \nabla \cdot \bar{U} \tag{7.4.29}
\end{equation*}
$$

In the first term in Eq. (7.4.28), $\bar{U}=\bar{V}_{C}$. This term is independent of $\bar{\xi}$, so $\nabla \cdot \bar{V}_{C}=0$. For the second term, $\bar{U}=\bar{\Omega} \times \bar{\xi}$. The same vector identity that led to vanishing of the monopole amplitude states that $\nabla \cdot(\bar{\Omega} \times \bar{\xi})$ is zero. Thus, we have reduced the dipole moment components to

$$
\begin{align*}
D_{j}= & -\frac{1}{4 \pi} i k \rho_{0} c \mathcal{V}\left(\bar{V}_{C} \cdot \bar{e}_{j}\right)-\frac{1}{4 \pi} i k \rho_{0} c \iiint_{\mathcal{V}}(\bar{\Omega} \times \bar{\xi}) \cdot \bar{e}_{j} d \mathcal{V} \\
& -\frac{1}{4 \pi} \iint_{\mathcal{S}} P(\bar{\xi}) n_{j}(\bar{\xi}) d \mathcal{S} \tag{7.4.30}
\end{align*}
$$

The fact that we have designated point $C$ to be the centroid of the body means that the first moment of position is zero, that is,

$$
\begin{equation*}
\iiint_{\mathcal{V}} \xi_{j} d \mathcal{V}=0 \tag{7.4.31}
\end{equation*}
$$

Thus, the middle integral vanishes. The dipole moment vector that results from adding the individual components of the remaining terms is

$$
\begin{equation*}
\bar{D}=-\frac{1}{4 \pi}\left[\rho_{0} \mathcal{V}\left(i \omega \bar{V}_{C}\right)+\iint_{\mathcal{S}} P(\bar{\xi}) \bar{n}(\bar{\xi}) d \mathcal{S}\right] \tag{7.4.32}
\end{equation*}
$$

To interpret the dipole moment, we observe that $\rho_{0} \mathcal{V}$ is the mass $m_{\text {disp }}$ of the body of fluid displaced by the vibrating body, and the complex amplitude of the acceleration of point $C$ is $i \omega \bar{V}_{C}$. Hence, the first term in the bracket is the force that would be required to impart to the displaced fluid the acceleration of the vibrating body. The second term also has a familiar explanation. Recall that $\bar{n}(\bar{\xi})$ is the normal to the vibrating body pointing into the fluid. The normal force exerted by the fluid on a surface patch $d \mathcal{S}$ of the vibrating body is $-\bar{n}(\bar{\zeta}) P(\bar{\xi}) d \mathcal{S}$. The resultant pressure force $\bar{F}_{p}$ exerted by the fluid is the surface integral of this term,

$$
\begin{equation*}
\bar{F}_{P}=-\iint_{\mathcal{S}} P(\bar{\xi}) \bar{n}(\bar{\xi}) d \mathcal{S} \tag{7.4.33}
\end{equation*}
$$

It follows that the dipole moment in Eq. (7.4.32) may be written as

$$
\begin{equation*}
\bar{D}=-\frac{1}{4 \pi}\left(m_{\mathrm{disp}} i \omega \bar{V}_{C}-\bar{F}_{P}\right) \tag{7.4.34}
\end{equation*}
$$

The force exerted on the fluid by the body is the reaction, that is, it is $-\bar{F}_{p}$. Thus, the dipole moment is $-1 / 4 \pi$ times the sum of the force required to accelerate the displaced fluid mass and the resultant force exerted on the fluid by the vibrating body.

To make this relation useful, it is necessary to quantify $\bar{F}_{P}$. Doing so requires an analysis specific to the vibrating body's shape. Nevertheless, certain aspects can be anticipated. For low frequencies, the pressure should be proportional to the body's acceleration, but opposed to it. Therefore, $\bar{F}_{p}$ must be proportional to $-i \omega \bar{V}_{C}$. Given that the first term in the dipole moment is proportional to the displaced mass, it is reasonable to anticipate that $\bar{F}_{P}$ is proportional to the density of the fluid. An aspect that might not be apparent is that $\bar{F}_{P}$ is not necessarily parallel to the acceleration. A vibrating thin disk exemplifies this behavior because the resultant of the pressure distribution must be perpendicular to the flat surface of the disk regardless of how the disk is oriented.

The corollary of the possible misalignment of $\bar{F}_{P}$ and $i \omega \bar{V}_{C}$ is that the proportionality is tensorial in nature. To describe it, we shall switch to a matrix description.

Let $\left\{F_{P}\right\}$ and $\left\{V_{C}\right\}$ be column vectors that are populated by the components of the respective quantities. The preceding general observations lead to definition of the factor of proportionality as a square matrix [ $W$ ], such that

$$
\begin{equation*}
\left\{F_{P}\right\}=-i \omega \rho_{0}[W]\left\{V_{C}\right\} \tag{7.4.35}
\end{equation*}
$$

The meaning of relation becomes somewhat clearer if we represent [ $W$ ] in partitioned form as a column vector, that is

$$
\left\{\begin{array}{l}
F_{x}  \tag{7.4.36}\\
F_{y} \\
F_{z}
\end{array}\right\}=-i \omega \rho_{0}\left[\begin{array}{l}
\left\{W^{(x)}\right\}^{\mathrm{T}} \\
\left\{W^{(y)}\right\}^{\mathrm{T}} \\
\left\{W^{(z)}\right\}^{\mathrm{T}}
\end{array}\right]\left\{V_{C}\right\}
$$

This representation shows that the components of $\bar{F}_{P}$ in the $x, y$, and $z$ directions are, respectively, the projection of $\bar{V}_{C}$ onto vectors $\bar{W}^{(x)}, \bar{W}^{(y)}$, and $\bar{W}^{(z)}$. Only if [ $W$ ] is proportional to the identity matrix will $\bar{F}_{P}$ be parallel to $\bar{V}_{C}$. Dimensional consistency requires that the elements of [ $W$ ] be volume quantities. Newton's second law states that $\bar{F}_{e}+\bar{F}_{P}=M(i \omega \bar{V})$, where $\bar{F}_{e}$ is the external force causing the body to move and $M$ is the body's mass. Substitution for $\bar{F}_{p}$ gives

$$
\begin{equation*}
\left\{F_{e}\right\}=i \omega\left[M[I]+\rho_{0}[W]\right]\left\{V_{C}\right\} \tag{7.4.37}
\end{equation*}
$$

This expression leads to $\rho_{0}[W]$ being referred to as the virtual mass, or the added mass.

Although determination of [ $W$ ] requires further analysis specific to the body's shape, certain properties are general. The same vector $\bar{F}_{p}$ must result from a specified vector $\bar{V}_{C}$, regardless of the orientation of the coordinate system associated with the respective components. This condition requires that [ $W$ ] be symmetric. A special property applies if the body is axisymmetric. Let this axis be $z$. The resultant force then must lie in the plane formed by $\bar{V}_{c}$ and $\bar{e}_{z}$ because the differential surface pressure resultant $-P(\bar{\xi}) \bar{n}(\bar{\xi}) d \mathcal{S}$ on points on either side of this plane balance. In addition, the force component transverse to $z$ must be invariant as the body is rotated about that axis with $\bar{V}_{C}$ fixed. These conditions require that [ $W$ ] be diagonal, with $W_{1,1,}=W_{2,2}$. Two shapes of interest are a sphere and a thin disk, ${ }^{15}$ for which the nonzero elements of the respective matrices are

$$
\begin{equation*}
\left(W_{\text {sphere }}\right)_{j, j}=\frac{2}{3} \pi a^{3}, \quad j=1,2,3 ; \quad\left(W_{\text {disk }}\right)_{3,3}=\frac{8}{3} a^{3} \tag{7.4.38}
\end{equation*}
$$

Determination of [ $W$ ] allows us to describe the dipole moment in matrix form. The relation is

$$
\begin{equation*}
\{D\}=-\frac{i \omega \rho_{0}}{4 \pi}[\mathcal{V}[I]+[W]]\left\{V_{c}\right\} \tag{7.4.39}
\end{equation*}
$$

[^12]Knowledge that the monopole part of the radiated field is zero, combined with the ability to evaluate the dipole contribution, usually is adequate to determine the pressure radiated by a rigid body vibrating at low frequencies. This is so because the dipole component will dominate the higher order terms. However, in some cases, it will be found that $\bar{D}=\overline{0}$. This situation will occur if the body oscillates in a pure rotation about the centroid $C$, so that $\bar{V}_{C}$ is zero. The pressure resultant is zero in such a motion because some pairs of points on the body's surface move in opposite sense as a result of the rotation. Consequently, the surface pressure distribution shows a $180^{\circ}$ change of phase, with the net result that different regions cancel. In other words, the small $k a$ field radiated by a body that executes a pure rotation will be dominated by various quadrupoles.

### 7.5 Numerical Methods for Radiation from Arbitrary Objects

The techniques we have developed are limited to specific configurations. Determination of the field radiated by a body that has an arbitrary shape requires a numerical simulation. We will examine three approaches having fundamentally different foundations. The first is the source superposition method, which is based on the notion that the radiating body may be replaced by a set of point sources in free space. Another approach converts the Kirchhoff-Helmholtz integral theorem to an integral equation for the surface pressure. A numerical analysis of this equation leads to the boundary element method. Both of these approaches are founded on acoustical analyses we have already carried out. The third is the finite element method, which is a general tool for solving field equations associated with the various media. We will see that each method has positive aspects and liabilities. The objective is to familiarize the reader with the available alternatives. Hence, our investigations will focus on the fundamental issues, rather than the details of their implementation. Each approach has several variants, but we will not delve into them.

### 7.5.1 Source Superposition

The concept here is quite simple-replace the radiating body with a collection of point sources. The idea is quite natural, but it leads to questions of how many sources are needed, where should they be placed, and how strong should each source be? Answering these questions is the contribution of the work by Koopman, Song, and Fahnline. ${ }^{16}$ They began by considering a continuous sheet of sources distributed over a surface interior to the radiating body. This was done to prove that the method is equivalent to the Kirchhoff-Helmholtz integral theorem. That is the junction at which we begin.

[^13]Fig. 7.19 Sources interior to a body whose normal component of surface velocity is $V_{S}\left(\bar{x}_{s}\right)$


Figure 7.19 depicts a body surrounded by an ideal fluid. Its surface $\mathcal{S}$ executes a specified vibration in which the velocity normal to the surface, positive into the fluid, is $\operatorname{Re}\left(V_{s}\left(\bar{x}_{s}\right) e^{i \omega t}\right)$. Distributed throughout the region surrounded by $\mathcal{S}$ is a set of $N$ point sources whose strengths will be determined. The objective is to determine the pressure at field point $\bar{x}_{0}$.

The source superposition method may be used to analyze two- and threedimensional situations. For the two-dimensional case, $G\left(\bar{x}, \bar{x}_{n}\right)$ is described by Eq. (7.3.47), and Eq. (6.4.31) describes the three-dimensional version. The pressure is the superposition of the contribution of each source. The strength of that contribution is the complex amplitude of the mass acceleration, which is $i \omega \rho_{0} \hat{Q}_{n}$, with $\hat{Q}_{n}$ being the volume velocity amplitude. Thus, the pressure at any point $\bar{x}$ is given by

$$
\begin{equation*}
P(\bar{x})=\sum_{n=1}^{N} i \omega \rho_{0} \hat{Q}_{n} G\left(\bar{x}, \bar{x}_{n}\right) \tag{7.5.1}
\end{equation*}
$$

An expression for the particle velocity is obtained by applying Euler's equation to the preceding. The gradient is evaluated at $\bar{x}$, and the free-space Green's function only depends on the distance between $\bar{x}_{n}$ and $\bar{x}$, so that

$$
\begin{equation*}
\bar{V}(\bar{x})=-\sum_{n=1}^{N} \hat{Q}_{n}\left(\frac{\bar{x}-\bar{x}_{n}}{r_{n}}\right) \frac{d}{d r_{n}} G\left(\bar{x}, \bar{x}_{n}\right), \quad r_{n}=\left|\bar{x}-\bar{x}_{n}\right| \tag{7.5.2}
\end{equation*}
$$

To determine the unknown volume velocities, we require that the particle velocity matches the known normal velocity distribution $V_{s}\left(\bar{x}_{s}\right)$. We cannot do so at every point on the surface, so we select a set of surface points, designated $\bar{\xi}_{m}$, at which this condition will be imposed. Let $M$ be the number of surface points selected for this purpose. Then, we have

$$
\begin{equation*}
\sum_{n=1}^{N} \hat{Q}_{n} \bar{n}\left(\bar{\xi}_{m}\right) \cdot\left(\frac{\bar{\xi}_{m}-\bar{x}_{n}}{r_{n}}\right) \frac{d}{d r_{n}} G\left(\bar{\xi}_{m}, \bar{x}_{n}\right)=-V_{s}\left(\bar{\xi}_{m}\right), m=1,2, \ldots, M \tag{7.5.3}
\end{equation*}
$$

The only quantities that are not set are the volume velocities. Thus, we have developed $M$ equations for these parameters.

Solution of these equations is expedited by writing them in matrix form. The volume velocities are arranged sequentially in a column vector. The coefficient of $\hat{Q}_{n}$ in equation $\# m$ is an element of the $M \times N$ rectangular array [ $U$ ],

$$
\begin{equation*}
U_{m, n}=\bar{n}\left(\bar{\xi}_{m}\right) \cdot\left(\frac{\bar{\xi}_{m}-\bar{x}_{n}}{r_{n}}\right) \frac{d}{d r_{n}} G\left(\bar{\xi}_{m}, \bar{x}_{n}\right) \tag{7.5.4}
\end{equation*}
$$

Thus, the task is to solve

$$
\begin{equation*}
[U]\{\hat{Q}\}=-\left\{V_{s}\right\} \tag{7.5.5}
\end{equation*}
$$

At this juncture, selection of the $\bar{x}_{n}$ locations for the sources and of the $\bar{\xi}_{m}$ locations for velocity matching enters the picture. If the number $M$ of the latter is less than $N$, then there are more sources than the number of available equations. Thus, it is necessary that $M$ at least equals $N$. Koopman, Song, and Fahnline set $M=N$, but this choice is problematic. If we were to select a different set of points to perform the matching, then it is likely that the $\hat{Q}_{n}$ values would be different. However, based on the proof that the source superposition and KHIT formalisms are equivalent, we expect the volume velocities to be unique. Thus, setting $M=N$ and increasing that number should eventually lead to a convergent result.

An alternative that has been followed matches the surface velocity at more points than the number of sources, $M>N$. Suppose we have identified a tentative solution for $\{\hat{Q}\}$ using $M>N$. Because there are more equations than the number of unknowns, it is unlikely that $\{\hat{Q}\}$ will satisfy all equations. A velocity matching error may be formed by substituting the solution into Eq. (7.5.5), then comparing the two sides of the equation. This gives an error residual $\{E\}$ whose definition is

$$
\begin{equation*}
\{E\}=[U]\{\hat{Q}\}+\left\{V_{s}\right\} \tag{7.5.6}
\end{equation*}
$$

A single error metric $E^{2}$ is the summed magnitude squared of all elements of $\{E\}$. This quantity is useful because it cannot be negative, so minimizing it assures that all elements of $\{E\}$ are a minimum. Another term for this metric is that it is the square of the Euclidean norm of $\{E\}$, which is obtained as a Hermitian dot product. The Hermitian of a matrix is the complex conjugate of its transpose, so the product is

$$
\begin{align*}
E^{2} \equiv \sum_{m=1}^{M}\left|E_{m}\right|^{2} \equiv\{E\}^{\mathrm{H}}\{E\}= & \{\hat{Q}\}^{\mathrm{H}}[U]^{\mathrm{H}}[U]\{\hat{Q}\}+\{\hat{Q}\}^{\mathrm{H}}[U]^{\mathrm{H}}\left\{V_{s}\right\}  \tag{7.5.7}\\
& +\left\{V_{s}\right\}^{\mathrm{H}}[U]\{\hat{Q}\}+\left\{V_{s}\right\}^{\mathrm{H}}\left\{V_{s}\right\}
\end{align*}
$$

We wish to find the set of source strengths that minimize $E^{2}$, which means that the value of $E^{2}$ must be stationary with respect to each $\hat{Q}_{j}$. References usually only
discuss the case where all quantities are real. The case where these quantities are complex is slightly more complicated, but the result is only different by replacing transpose operations by Hermitians, that is,

$$
\begin{equation*}
[U]^{\mathrm{H}}[U]\{\hat{Q}\}=-[U]^{\mathrm{H}}\left\{V_{s}\right\} \tag{7.5.8}
\end{equation*}
$$

This formulation is said to be the method of linear least squares. The coefficient matrix [ $U$ ] has $M$ rows (the number of velocity equations) and $N$ columns (the number of sources). Thus, $[U]^{\mathrm{H}}[U]$ is $N \times N$. The inverse of this matrix is called the Morse-Penrose inverse, but any method may be used to solve for $\hat{Q}$. After those values have been obtained, the pressure at any field point may be determined from Eq. (7.5.1).

The formulation seems to be straightforward, but its implementation requires that several issues be addressed. Most significant is the possibility that $[U]^{\mathrm{H}}[U]$ is illconditioned. As explained by Ochmann, ${ }^{17}$ this situation might occur if the radiating body's shape is drastically different from spherical, as would be the case with a long cylinder. Another reason might be that the distribution of sources is inconsistent with the symmetry of the body. For example, a rectangular box should have sources placed equally on both sides of each of its midplanes. Ill-conditioned situations typically are addressed by application of singular value decomposition, ${ }^{18}$ but doing so requires greater computational effort.

The question still remains as to where the sources should be located. Koopman et al. provide some guidelines, as does Ochmann. ${ }^{19}$ Because the equations are quite easy to formulate, practitioners typically use a very large number of sources placed consistently with the symmetry of the system. For example, consider radiation from a cylinder whose surface vibration is axisymmetric. Many sources should be distributed along the cylinder's axis in order that the field has no variation around any circumferential circle. When this tactic is used, then the number of surface points $\bar{\xi}_{m}$ for matching the surface velocity would be chosen to be much greater than $N$. Some users distribute the source and surface locations uniformly, and some use a random number generator for that purpose.

Regardless of how one selects these locations, the result should be verified. One way of doing so is to redo the analysis with another set of source and surface points. Another way to verify the solution for $\hat{Q}_{n}$ values is to use Eq. (7.5.2) to evaluate $\bar{n}\left(\bar{x}_{s}\right) \cdot \bar{V}\left(\bar{x}_{s}\right)$ at several points on the surface that were not used to formulate Eq. (7.5.8). The degree to which this set of normal velocities differ from the known values of $V_{s}\left(\bar{x}_{s}\right)$ provides a measure of the error in the method.

[^14]
### 7.5.2 Boundary Element Method

A very different approach is founded on the KHIT. A preview of this method arose in Example 7.9, where the KHIT was used to evaluate the pressure at field points that are very close to the surface. We begin with the radiation version of KHIT, in which the field point $\bar{x}_{0}$ is exterior to the vibrating body. The fundamental equation we seek has the field point on the surface. This is not a trivial matter because the integrand of the KHIT is singular when the surface point $\bar{x}_{s}$ at which the integrand is evaluated coincides with $\bar{x}_{0}$.

## The Surface Helmholtz Integral Equation

The singularity of the KHIT integrand is handled by bringing the field point to the surface in a limiting process. In Fig. 7.20, the field point is at distance $\varepsilon$ from the surface measured in the direction of the surface normal that intersects this field point. The surface point for this normal is designated $\bar{x}_{0}^{\prime}$, so that the field point's position is $\bar{x}_{0}=\bar{x}_{0}^{\prime}+\varepsilon \bar{n}\left(\bar{x}_{0}^{\prime}\right)$.

Fig. 7.20 Placement of a field point close to a vibrating surface


The idea is to decompose the surface into two parts: $\mathcal{S}_{0}$ is a small region that surrounds $\bar{x}_{0}^{\prime}$ and $\mathcal{S}_{1}$ is the remainder of $\mathcal{S}$. Because $\bar{x}_{0}$ is situated in the fluid domain, we set $\chi=1$ in Eq.(7.4.6), so the KHIT may be written as

$$
\begin{gather*}
P\left(\bar{x}_{0}\right)=\mathcal{I}_{0}+\mathcal{I}_{1} \\
\mathcal{I}_{j}=\iint_{\mathcal{S}_{j}}\left[i \omega \rho_{0} \bar{n}\left(\bar{x}_{s}\right) \cdot V\left(\bar{x}_{s}\right) G\left(\bar{x}_{0}, \bar{x}_{s}\right)+P\left(\bar{x}_{s}\right) \bar{n}\left(\bar{x}_{s}\right) \cdot \nabla G\left(\bar{x}_{0}, \bar{x}_{s}\right)\right] d \mathcal{S} \tag{7.5.9}
\end{gather*}
$$

In order for $\bar{n}\left(\bar{x}_{0}^{\prime}\right)$ to be a unique normal vector, it must be that $\mathcal{S}$ is smooth at that location, that is, there must be a unique tangent plane there. (This excludes sharp corners, which we will address after the basic derivation.) When the geometry fits this specification, we may consider $\mathcal{S}_{0}$ to be a small circle of radius $\Delta$ lying in the tangent plane and centered on $\bar{x}_{0}^{\prime}$. This is the configuration depicted in Fig. 7.21.

The procedure we will follow is to consider the limiting behavior of the integral over $\mathcal{S}_{0}$ as $\varepsilon \rightarrow 0$ with $\Delta$ held at a small but finite value. After that limit is established,

Fig. 7.21 Definition of a spherical coordinate system for evaluating the principal part of the KHIT when the field point is situated on the vibrating surface

we will take the limit as $\Delta \rightarrow 0$. The resulting expression for the integral over $\mathcal{S}_{0}$ is called the Cauchy principal part of the integral over $\mathcal{S}$. We will derive the result for a three-dimensional system, so the free-space Green's function will be $G\left(\bar{x}_{0}, \bar{x}_{s}\right)=e^{-i k r} /(4 \pi r)$.

The geometric properties of the terms in the integral over $\mathcal{S}_{0}$ are described in term of a spherical coordinate system centered on $\bar{x}_{0}$. The radius $\Delta$ is small, and it eventually will be reduced to zero, which means that the surface pressure and normal velocity may be taken to be constant over $\mathcal{S}_{0}$ at their values at the center point. Thus, the portion of KHIT associated with this region is

$$
\begin{align*}
\mathcal{I}_{0} & \approx i \rho_{0} \omega V_{s}\left(\bar{x}_{0}^{\prime}\right) \iint_{\mathcal{S}_{0}} G\left(\bar{x}_{0}, \bar{x}_{s}\right) d \mathcal{S}+P\left(\bar{x}_{0}^{\prime}\right) \iint_{\mathcal{S}_{0}} \bar{n}\left(\bar{x}_{s}\right) \cdot \nabla G\left(\bar{x}_{0}, \bar{x}_{s}\right) d \mathcal{S} \\
& =\frac{1}{4 \pi} i \rho_{0} \omega V_{s}\left(\bar{x}_{0}^{\prime}\right) \iint_{\mathcal{S}_{0}} \frac{e^{-i k r}}{r} d \mathcal{S}+\frac{1}{4 \pi} P\left(\bar{x}_{0}^{\prime}\right) \iint_{\mathcal{S}_{0}} \bar{e}_{z} \cdot\left(\frac{\bar{x}_{s}-\bar{x}_{0}}{r}\right) \frac{d}{d r}\left(\frac{e^{-i k r}}{r}\right) d \mathcal{S} \tag{7.5.10}
\end{align*}
$$

Each of the terms in the integrand is readily described in terms of the spherical coordinate system in Fig. 7.21. The axis of this coordinate system is normal to the circle, so the geometry is asymmetric. Consequently, we may use a ring element to describe the differential area, so that $d \mathcal{S}=2 \pi R d R$. The value of $R$ is related to the radial distance by $R^{2}=r^{2}-\varepsilon^{2}$. The differential of this relation is $R d R=r d r$ because $\varepsilon$ is the constant distance to the surface. The range of radial distances covered by $\mathcal{S}_{0}$ is $\varepsilon \leq r \leq\left(\varepsilon^{2}+\Delta^{2}\right)^{1 / 2}$. This change of variables converts the integral to

$$
\begin{align*}
\mathcal{I}_{0}= & \frac{1}{2} i \rho_{0} \omega V_{s}\left(\bar{x}_{0}^{\prime}\right) \int_{\varepsilon}^{\left(\varepsilon^{2}+\Delta^{2}\right)^{1 / 2}} e^{-i k r} d r \\
& +\frac{1}{2} P\left(\bar{x}_{0}^{\prime}\right) \int_{\varepsilon}^{\left(\varepsilon^{2}+\Delta^{2}\right)^{1 / 2}} \bar{e}_{z} \cdot\left(\frac{\bar{x}_{s}-\bar{x}_{0}}{r}\right)\left(-i k-\frac{1}{r}\right) e^{-i k r} d r \tag{7.5.11}
\end{align*}
$$

The dot product in the second integral is expressible in terms of $\varepsilon$ and the radial distance, specifically,

$$
\begin{equation*}
\bar{e}_{z} \cdot \frac{\left(\bar{x}_{s}-\bar{x}_{0}\right)}{r}=-\cos \psi=-\frac{\varepsilon}{r} \tag{7.5.12}
\end{equation*}
$$

In addition, the smallness of $\varepsilon$ and $\Delta$ leads to simplification of both integrals, because we may set $\exp (-i k r)=1$ and approximate the last factor in the second integrand as $-1 / r$. The result is that the integral over $\mathcal{S}_{0}$ reduces to

$$
\begin{align*}
\mathcal{I}_{0} & =\frac{1}{2} i \rho_{0} \omega V_{s}\left(\bar{x}_{0}^{\prime}\right) \int_{\varepsilon}^{\left(\varepsilon^{2}+\Delta^{2}\right)^{1 / 2}} d r+\frac{1}{2} P\left(\bar{x}_{0}^{\prime}\right) \int_{\varepsilon}^{\left(\varepsilon^{2}+\Delta^{2}\right)^{1 / 2}} \frac{\varepsilon}{r^{2}} d r \\
& =\frac{1}{2} i \rho_{0} \omega V_{s}\left(\bar{x}_{0}^{\prime}\right)\left[\left(\varepsilon^{2}+\Delta^{2}\right)^{1 / 2}-\varepsilon\right]+\frac{1}{2} P\left(\bar{x}_{0}^{\prime}\right) \varepsilon\left[\frac{1}{\varepsilon}-\frac{1}{\left(\varepsilon^{2}+\Delta^{2}\right)^{1 / 2}}\right] \tag{7.5.13}
\end{align*}
$$

The limit of $\mathcal{I}_{0}$ as $\varepsilon \rightarrow 0$ is

$$
\begin{equation*}
\mathcal{I}_{0}=\frac{1}{2} i \rho_{0} \omega V_{s}\left(\bar{x}_{0}^{\prime}\right) \Delta+\frac{1}{2} P\left(\bar{x}_{0}^{\prime}\right) \tag{7.5.14}
\end{equation*}
$$

When we take the limit as $\Delta \rightarrow 0$, the first term vanishes but the second is unaltered. In regard to the integral over $\mathcal{S}_{1}$, we recognize that it approaches the original integral over the entire surface $\mathcal{S}$, except that the singular point, $\bar{x}_{s}=\bar{x}_{0}^{\prime}$, is excluded from the domain. (How this is done in a computational scheme will be discussed.) Thus, to evaluate Eq.(7.5.9), we carry out a regular integration over the entire surface $S$ with $\bar{x}_{0}$ excluded, then add $P\left(\bar{x}_{0}^{\prime}\right) / 2$. When we bring the latter term to the left side, we find that $P\left(\bar{x}_{0}^{\prime}\right) / 2$ equals the regular integral.

Let us step back to take an overview of the KHIT before we write the final form. The left side is $\chi P\left(\bar{x}_{0}\right)$. Previously, we found that $\chi=1$ if $\bar{x}_{0}$ is outside the radiating body, and $\chi=0$ if $\bar{x}_{0}$ is inside the radiating body. Here, we brought $\bar{x}_{0}$ to the surface. Doing so led to $\chi=1 / 2$. This result was derived by assuming that the surface is smooth. If $\bar{x}_{0}$ is situated at a corner, then the smoothness condition is not met. That situation is addressed by the concept of a solid angle. Figure 7.22 shows the four cases for $\bar{x}_{0}$ : (a) exterior, (b) interior, (c) smooth surface, and (d) corner. For each, a small sphere of radius $\delta$ is centered on $\bar{x}_{0}$.

The portion of the surface area of the sphere that lies in the fluid domain is defined to be $\Gamma \delta^{2}$, where $\Gamma$ is the solid angle. Thus, $\Gamma=4 \pi$ for case (a), $\Gamma=0$ for case (b),


Fig. 7.22 Parameter $\chi$ for the KHIT is obtained by surrounding the field point by a small sphere of radius $\delta$. The solid angle $\delta^{2} \Gamma$ is the portion of the sphere's surface area that is situated in the fluid
$\Gamma=2 \pi$ for case (c), and $\Gamma$ is between $2 \pi$ and $4 \pi$ for case (d) because more than half the sphere is in the fluid at that corner. The value of $\Gamma$ at a corner depends on whether the slope is discontinuous in one direction or two orthogonal directions. For example, if $\bar{x}_{0}$ is at the edge of cylinder with flat ends, then $\Gamma=3 \pi$, whereas $\Gamma=(7 / 8) 4 \pi$ if $\bar{x}_{0}$ is at the corner of a rectangular box. The parameter $\chi\left(\bar{x}_{0}\right)$ describes the fraction of a full sphere that lies in the fluid. In general, it is

$$
\begin{equation*}
\chi\left(\bar{x}_{0}\right)=\frac{\Gamma}{4 \pi} \tag{7.5.15}
\end{equation*}
$$

An extremely useful feature stems from the fact that KHIT applies to acoustic cavities, as well as radiation problems. The sole modifications required to treat either situation is that $\bar{n}$ must point into the fluid and the surface normal velocity must be defined to be positive if it is in the sense of $\bar{n}$. It follows that a similar modification of the SHIE will make it applicable to cavities. In that case, the solid angle would be defined by the extent of the interior fluid surrounding $\bar{x}_{0}$.

The KHIT for all cases is described by

$$
\begin{align*}
\chi\left(\bar{x}_{0}\right) P\left(\bar{x}_{0}\right) & =\frac{1}{4 \pi} \iint_{\mathcal{S}}\left[i \omega \rho_{0} V_{s}\left(\bar{x}_{s}\right)-P\left(\bar{x}_{s}\right) \bar{n}\left(\bar{x}_{s}\right) \cdot\left(\frac{\bar{x}_{s}-\bar{x}_{0}}{\left|\bar{x}_{n}-\bar{x}_{s}\right|}\right)\right.  \tag{7.5.16}\\
& \left.\times\left(\frac{i k\left|\bar{x}_{n}-\bar{x}_{s}\right|+1}{\left|\bar{x}_{n}-\bar{x}_{s}\right|}\right)\right] \frac{e^{-i k\left|\bar{x}_{n}-\bar{x}_{s}\right|}}{\left|\bar{x}_{n}-\bar{x}_{s}\right|} d \mathcal{S}
\end{align*}
$$

It is implicit to this expression that $\bar{x}_{s}=\bar{x}_{0}$ is excluded from the integration domain if $\bar{x}_{0}$ is on the surface. In that case, the equation is called the surface Helmholtz integral equation. We will use the abbreviation SHIE to refer to it.

## Discretized Implementation

In a radiation problem, we seek the pressure field generated by a specified velocity distribution on the vibrating surface. The SHIE states that the pressure on the surface depends on the properties of the pressure on the surface. Just as an integral is sometimes called an antiderivative, such a relation could be considered to be an "antidifferential equation." (Its proper description is that is Fredholm integral equation of the second kind.) It often is necessary to use numerical methods to solve a differential equation, and that is the situation for the SHIE. This process begins by dividing the surface into a set of small patches, which we will denote as $\mathcal{S}_{j}$. Each patch is one element situated on the surface, from which the description as a boundary element formulation follows.

The integral over a patch depends on the geometrical properties of the surface and the spatial distribution of $P\left(\bar{x}_{s}\right)$ and $V_{s}\left(\bar{x}_{s}\right)$. Some formulations have used interpolating functions to describe these properties. We will adopt a simpler approach in which both fields are considered to be constant over a patch. This is the formulation that was implemented to develop CHIEF, ${ }^{20}$ which was one of the first effective numerical

[^15]codes given wide distribution. The meaning of the abbreviation will emerge in the course of the development.

Let $\bar{x}_{j}$ denote the location of the center of element $\mathcal{S}_{j}$, and define $P_{j} \equiv P\left(\bar{x}_{j}\right)$ and $V_{j} \equiv V_{s}\left(\bar{x}_{j}\right)$. Because $\bar{x}_{0}$ in the SHIE is situated on the surface, $\bar{x}_{0}$ is taken to be one of these center points. The pressure and velocity are factored out of the integral over a patch because the requisite smallness of all $\mathcal{S}_{j}$ makes negligible their variation over that region, Thus, Eq. (7.5.16) becomes

$$
\begin{align*}
4 \pi \chi_{n} P_{n}= & \sum_{j=1}^{N} i \omega \rho_{0} V_{j} \iint_{\mathcal{S}_{j}} \frac{e^{-i k\left|\bar{x}_{n}-\bar{x}_{s}\right|}}{\left|\bar{x}_{n}-\bar{x}_{s}\right|} d \mathcal{S} \\
& -\sum_{j=1}^{N} P_{j} \iint_{\mathcal{S}_{j}} \bar{n}\left(\bar{x}_{s}\right) \cdot\left(\bar{x}_{s}-\bar{x}_{n}\right)\left(\frac{i k\left|\bar{x}_{n}-\bar{x}_{s}\right|+1}{\left|\bar{x}_{n}-\bar{x}_{s}\right|^{3}}\right) e^{-i k\left|\bar{x}_{n}-\bar{x}_{s}\right|} d \mathcal{S} \tag{7.5.17}
\end{align*}
$$

Both integrands depend only on the geometrical properties of $\mathcal{S}$ and the selection of the points on $\mathcal{S}$. Thus, the integrals constitute a set of coefficients that are independent of the pressure and surface velocity,

$$
\begin{gather*}
B_{n, j}=\iint_{\mathcal{S}_{j}} \bar{n}\left(\bar{x}_{s}\right) \cdot\left(\bar{x}_{s}-\bar{x}_{n}\right)\left(\frac{i k\left|\bar{x}_{n}-\bar{x}_{s}\right|+1}{\left|\bar{x}_{n}-\bar{x}_{s}\right|^{3}}\right) e^{-i k\left|\bar{x}_{n}-\bar{x}_{s}\right|} d \mathcal{S} \\
C_{n, j}=\iint_{\mathcal{S}_{j}} \frac{e^{-i k\left|\bar{x}_{n}-\bar{x}_{s}\right|}}{\left|\bar{x}_{n}-\bar{x}_{s}\right|} d \mathcal{S} \tag{7.5.18}
\end{gather*}
$$

These coefficients are evaluated by two different schemes. No singularity arises in the off-diagonal terms, $j \neq n$, because $\bar{x}_{n}$ is not situated in $\mathcal{S}_{j}$. These terms are evaluated by a two-dimensional Gaussian integration scheme, which uses the value of the integrand at points interior to $\mathcal{S}_{j}$. A low-order scheme using four points is illustrated in Fig. 7.23, but some investigations ${ }^{21}$ have suggested that more points are required for accurate results.


Fig. 7.23 A four-point Gaussian integration rule for integrating over surface patch $\mathcal{S}_{j}$

[^16]Evaluation of the diagonal coefficients $B_{n, n}$ and $C_{n, n}$ recognizes that the Cauchy principal part gives the contribution of the first integral at $\bar{x}_{s}=\bar{x}_{n}$. The numerical integration scheme excludes this point by further dividing the patch in Fig. 7.23 into four quadrants indicated there by dashed lines. A sixteen-point Gaussian integration rule is used for each subpatch. Because the Gaussian points are internal to each subregion, evaluation of the integrands at $\bar{x}_{n}$ is excluded.

The discretized form of the SHIE, Eq. (7.5.17), may be written in matrix form as

$$
\begin{equation*}
[4 \pi[\chi]+[B]]\{P\}=[C]\{V\} \tag{7.5.19}
\end{equation*}
$$

where $[\chi]$ is a diagonal array of the value of $\chi$ at each center point $\bar{x}_{n}$; so $\chi_{n, n}=1 / 2$ except for points at corners and edges. Both $[B]$ and $[C]$ are $N \times N$ complex arrays, which is the number of unknown $P\left(x_{n}\right)$. It might seem that solving the equations for the $\{P\}$ given $\{V\}$ is a straightforward task. However, this is only true if the domain is an interior cavity. For the exterior radiation problem, there are certain frequencies at which difficulties arise.

All cavities are like closed one-dimensional waveguides and spherical cavities, in the sense that an infinite number of modes exist if the boundary is rigid. A mode is a pressure field that can exists without excitation; the frequency of that field is the natural frequency. Why is this important to the radiation problem? The KHIT for interior and exterior domains have similar appearance. It was proven by Schenck ${ }^{22}$ that if the frequency matches a natural frequency of the domain inside the vibrating surface, then Eq. (7.5.19) has no unique solution. Some individuals refer to this condition as a "forbidden internal cavity resonance." If the frequency for a specific computation was exactly one of these natural frequencies, the matrix $[4 \pi[\chi]+[B]]$ would be rank-deficient, so its inverse would not exist. The more common occurrence is that the overall computation entails a frequency sweep, and one or more of the frequencies in the swept range is close to a natural frequency for the enclosed region. Then, the matrix will be ill-conditioned. If one proceeds to solve for $\{P\}$ corresponding to this matrix, numerical errors will be greatly magnified, making the solution unusable.

How to proceed in this situation depends on what one knows about the modes of the domain contained inside $\mathcal{S}$. In some special cases, like a sphere or a flatended circular cylinder, the modal solution can be obtained analytically. In that case, one could simply omit any results obtained from Eq. (7.5.19) for frequencies close to natural frequencies. However, for arbitrary shapes, we seldom know what the interior natural frequencies are.

Loss of uniqueness at the forbidden frequencies has been addressed in several ways, each resulting in a different boundary element computer code. CHIEF apparently was the first general procedure. It is based on a very simple idea. The approximation of the SHIE in Eq. (7.5.19) has reduced rank at a forbidden frequency, which means that the number of independent equations is less than the size of $\{P\}$. Extra equations are obtained by using the original Helmholtz integral equation, Eq. (7.4.6),

[^17]for the case where the fluid is exterior to the vibrating surface and the field point $\bar{x}_{0}$ is interior to it. The field point is not within the fluids, which means that $\chi=0$, so the integral should evaluate to zero. As is done for the surface points, the values of $P$ and $V_{s}$ are factored out of the integral over each $\mathcal{S}_{j}$, and Gaussian integration is used to evaluate the contribution of the geometric terms in each integrand. Because $\bar{x}_{0}$ is situated inside $\mathcal{S}$, the integrand is never singular, so the Gaussian integration scheme used to evaluate the off-diagonal terms of $[B]$ and $[C]$ is used here also. Each interior field point to which this analysis we applied leads to an extra equation to supplement what has been lost. The computer code name "CHIEF", which is an acronym for Combined Helmholtz Integral Equation Formulation, was the first to implement this concept.

It is known that the rank of $[B]$ is less than $N$ at a forbidden frequency, but the actual rank is not known. Without that knowledge, it is not known how many equations for interior points are required to make the system of equations solvable. Furthermore, as was noted, the interior field is a mode at a forbidden frequency. Such modes have nodal surfaces, along which the pressure is zero. The Helmholtz integral equation will give zero identically if $\bar{x}_{0}$ is situated on a nodal surface, so it will not be an additional equation to supplement the equations for points on the surface. The strategy that addresses both issues uses far more interior points than the number that is required. Let $\bar{\xi}_{n}$ be the positions of these interior points, whose location may be selected in any convenient manner, including randomly. Let $M$ be the number of such points. Then, because $\chi=0$ for interior points, the numerical approximation of the Helmholtz integral equation at such points gives

$$
\begin{equation*}
\left[B^{\prime}\right]\{P\}=\left[C^{\prime}\right]\{V\} \tag{7.5.20}
\end{equation*}
$$

where the coefficient matrices are

$$
\begin{gather*}
B_{n, j}^{\prime}=\iint_{\mathcal{S}_{j}} \bar{n}\left(\bar{x}_{s}\right) \cdot\left(\bar{x}_{s}-\bar{\xi}_{n}\right)\left(\frac{i k\left|\bar{\xi}_{n}-\bar{x}_{s}\right|+1}{\left|\bar{\xi}_{n}-\bar{x}_{s}\right|^{3}}\right) e^{-i k\left|\bar{\xi}_{n}-\bar{x}_{s}\right|} d \mathcal{S}  \tag{7.5.21}\\
C_{n, j}=\iint_{\mathcal{S}_{j}} \frac{e^{-i k\left|\bar{\xi}_{n}-\bar{x}_{s}\right|}}{\left|\bar{\xi}_{n}-\bar{x}_{s}\right|} d \mathcal{S}
\end{gather*}
$$

The equations for the interior points are stacked below Eq. (7.5.19) for the surface points. There are $M$ of the former and $N$ of the latter, so the equations to be solved are

$$
\left[\begin{array}{c}
4 \pi[\chi]+[B]  \tag{7.5.22}\\
{\left[B^{\prime}\right]}
\end{array}\right]\{P\}=[C]\{V\}
$$

This represents a total of $N+M$ equations for the $M$ values contained in $\{P\}$. The equations are overdetermined, as they were for the source superposition method, Eq. (7.5.5). This condition is handled in the same way, specifically by invoking the method of linear least squares. The result is that the coefficient matrix is defined to be
an array [ $U$ ] whose size is $(N+M) \times N$, and the equation to be solved is multiplied by $[U]^{\mathrm{T}}$. This leads to

$$
\begin{gather*}
{[U]=\left[\begin{array}{c}
{[4 \pi[\chi]+[B]]} \\
{\left[B^{\prime}\right]}
\end{array}\right]}  \tag{7.5.23}\\
{[U]^{\mathrm{T}}[U]\{P\}=[U]^{\mathrm{T}}[C]\{V\}}
\end{gather*}
$$

After the solution for $\{P\}$ has been obtained, the pressure at any field point exterior to $\mathcal{S}$ is found by using the same formulation as that used for the interior points. If the field point is in the farfield, this evaluation would be based on the farfield form in Eq. (7.4.18).

CHIEF is just one of many boundary element codes. They differ in the way that the variation of $P\left(\bar{x}_{s}\right)$ and $V_{s}\left(\bar{x}_{s}\right)$ over a patch is described, as well as how the integral over the surface patches is obtained. In addition, some use a different integral equation to handle the forbidden frequency issue. One that has had some popularity is obtained by differentiating the Helmholtz integral equation to obtain an integral equation for $V_{s}$ on the surface. The combination of that equation and SHIE at the forbidden frequencies is called the Burton-Miller formulation. ${ }^{23}$

Boundary element formulations have some common attributes. One is that they tend to have a large number of variables. This is a consequence of a general guideline that the surface patches should be approximately square, with each side being no bigger than one-sixth of the acoustic wavelength at the frequency for the evaluation. The number of equations is increased by the supplemental equations required to handle the forbidden frequency issue. The requisite computation resources increase substantially as the frequency is increased in a sweep. This is so because one has the choice of setting the surface mesh to fit the one-sixth wavelength requirement at the highest frequency, or else rezoning the mesh as the frequency increases. Furthermore, all coefficient matrices are functions of frequency, so they must be recomputed at each step in the sweep. Another attribute that causes a boundary element model to require substantial computational resources is the fact that $[B]$ and $\left[B^{\prime}\right]$ are full matrices, so only standard solution algorithms may be employed. This is contrasted by the finite element method, whose matrices are diagonally dominant. Many algorithms that greatly reduce core memory requirements and the number of operations have been developed to solve such equations.

### 7.5.3 Finite Element Method

A variety of finite element formulations have been developed. We will consider a simple one in order to expose the broad concepts, and then discuss ways in which

[^18]it may be used to study radiation into an infinite domain. As with the preceding developments, we will only consider the frequency domain.

One can find finite element formulations for acoustics that employ displacement as a primary variable, either individually, or in combination with pressure. We will examine a simpler approach that only uses pressure. The derivation will apply the Galerkin method to the Helmholtz equation. The first step is to decompose the domain $\mathcal{V}$ of the fluid into small pieces, which are the finite elements. We shall denote a generic element as $\mathcal{V}_{n}$. Each element may have any shape; common ones are a box or a tetrahedron. In general, a set of mesh points is defined in each $\mathcal{V}_{n}$. Some of these points may be interior to the element. However, placing some on the surface of the element will assure continuity of pressure because those points will also lie on one or more adjacent elements. Interpolation functions, typically polynomials, are used to represent the pressure within an element.

To understand the basic aspects of the formulation, a one-dimensional waveguide is described in Fig. 7.24. The waveguide is segmented into elements that are numbered sequentially by index $n=1,2, \ldots N$. The length of element $n$ is $L_{n}$, which need not be the same for all elements. A mesh point is located at the end of each element, and no interior points are defined. A local coordinate $x_{n}$ is used to measure the axial position within element $n$.


Fig. 7.24 Numbering scheme for a finite element model of a one-dimensional waveguide

The mesh points within an element are numbered with the index $\alpha=1$ for the left node and $\alpha=2$ for the right node. A superscript will be used to indicate which element is under consideration if there is ambiguity. Therefore, the pressure at a node is denoted as $P_{\alpha}^{(n)}$. The lowest order interpolating function is linear. The corresponding description of the pressure distribution within an arbitrary element is

$$
P^{(n)}\left(x_{n}\right)=\left\{\begin{array}{l}
P_{1}^{(n)}\left(1-\frac{x_{n}}{L_{n}}\right)+P_{2}^{(n)} \frac{x_{n}}{L_{n}}, \quad 0 \leq x_{n} \leq L_{n}  \tag{7.5.24}\\
0 \text { otherwise }
\end{array}\right.
$$

The interpolating functions are denoted as $\mathcal{N}_{\alpha}^{(n)}\left(x_{n}\right)$, where $n$ is the element number and $\alpha$ is the mesh point index. Thus, the pressure within an element may be written as

$$
\begin{equation*}
P^{(n)}\left(x_{n}\right)=\sum_{\alpha=1}^{2} \mathcal{N}_{\alpha}^{(n)}\left(x_{n}\right) P_{\alpha}^{(n)} \tag{7.5.25}
\end{equation*}
$$

If one were to use higher order interpolating functions, the sole alteration in this expression would be adjustment of the summation range to match the commensurate increase in the number of mesh points per element. Because the interpolating functions for an element are defined to be zero outside their element, the pressure at any position may be considered to be a sum over all elements of the preceding expression, that is,

$$
\begin{equation*}
P=\sum_{n=1}^{N} \sum_{\alpha=1}^{2} \mathcal{N}_{\alpha}^{(n)}\left(x_{n}\right) P_{\alpha}^{(n)} \tag{7.5.26}
\end{equation*}
$$

This representation does not account for the fact that some mesh points are shared by elements, that is, $P_{2}^{(1)}=P_{1}^{(2)}, P_{2}^{(2)}=P_{1}^{(3)}, \ldots, P_{2}^{(N-1)}=P_{1}^{(N)}$. This overlap is described by assigning a unique index $\gamma$ to each mesh point, and defining $\hat{P}_{\gamma}$ to be the pressure at that point. These values are the global pressures, whereas the $P_{\alpha}^{(n)}$ are the local values. The latter may be obtained by using a filtering matrix $\left[S^{(n)}\right]^{\alpha}$ to pick out the appropriate elements of the global set $\{\hat{P}\}$,

$$
\left\{\begin{array}{l}
P_{1}^{(n)}  \tag{7.5.27}\\
P_{2}^{(n)}
\end{array}\right\}=\left[S^{(n)}\right]\{\hat{P}\}
$$

For example, the filtering matrices for the first two elements in Fig. 7.24 are

$$
\left[S^{(1)}\right]=\left[\begin{array}{lllll}
1 & 0 & 0 & 0 & \cdots  \tag{7.5.28}\\
0 & 1 & 0 & 0 & \cdots
\end{array}\right], \quad\left[S^{(2)}\right]=\left[\begin{array}{lllll}
0 & 1 & 0 & 0 & \cdots \\
0 & 0 & 1 & 0 & \cdots
\end{array}\right]
$$

The $\left[S_{n}\right]$ matrices are more commonly referred to as the connectivity matrix.
Although these relations have been derived for planar waves in a waveguide, they are applicable with a few modifications to a three-dimensional field. Position within an element is described by a local coordinate system $x_{n} y_{n} z_{n}$. Mesh points are distributed over an element, and more than two points per element are required. Other than the higher dimensionality, the basic relations are as above. The pressure at any position $\bar{x}_{n}$ within an element is a sum of terms consisting of an interpolating function multiplied by the associated mesh pressure. The matrix representation of this ansatz is

$$
\begin{equation*}
P\left(\bar{x}_{n}\right)=\left[\mathcal{N}^{(n)}\left(\bar{x}_{n}\right)\right]\left\{P^{(n)}\right\} \tag{7.5.29}
\end{equation*}
$$

This is the fundamental description of the pressure field. Only the mesh pressures are unknown. As before, the number of columns for the interpolating matrix equals the number of mesh points in an element, and it has one row.

The connectivity matrix $\left[S^{(n)}\right]$ still filters the pressures $\left\{P^{(n)}\right\}$ at an element's mesh points from the global set of pressures. The matrix representation of these relations is

$$
\begin{equation*}
\left\{P^{(n)}\right\}=\left[S^{(n)}\right]\{\hat{P}\} \tag{7.5.30}
\end{equation*}
$$

Of course, the number of rows for $\left[S^{(n)}\right]$ will be the number of mesh points in an element, which is greater than two. Equation (7.5.29) gives the pressure in a specific element, and the interpolating functions are defined to give zero if the position is outside that element. Thus, substitution of Eq. (7.5.30) into (7.5.29) gives the contribution of the mesh points within an element to the entire field. Adding these relations for all elements gives a representation of the pressure throughout the fluid,

$$
\begin{equation*}
P(\bar{x})=\sum_{n=1}^{N}\left[\mathcal{N}^{(n)}\left(x_{n}, y_{n}, z_{n}\right)\right]\left[S^{(n)}\right]\{\hat{P}\} \tag{7.5.31}
\end{equation*}
$$

Various rules for finite elements correspond to using different interpolating functions, and/or different shapes for the elements. The preceding will accommodate any rule.

The next task is to derive the equations governing the global pressures. Equation (7.5.31) will not satisfy the Helmholtz equation. The remainder after its substitution represents an error whose value will depend on the location within $\mathcal{V}$. The Galerkin method requires that this error be orthogonal to an arbitrary test function $\Psi$ that lies in the same functional space as the series for $P$. This means that the test function is like Eq. (7.5.31), except that its coefficients are an arbitrary set of values $\eta_{n}$. In other words

$$
\begin{equation*}
\Psi=\sum_{n=1}^{N}\left[\mathcal{N}^{(n)}\left(x_{n}, y_{n}, z_{n}\right)\right]\left[S^{(n)}\right]\{\eta\} \tag{7.5.32}
\end{equation*}
$$

The orthogonality condition is defined relative to an inner product of functions over the entire domain $\mathcal{V}$. Thus, the condition required in the Galerkin method is

$$
\begin{equation*}
\iiint_{\mathcal{V}}\left(\nabla^{2} P+k^{2} P\right) \Psi d \mathcal{V}=0 \tag{7.5.33}
\end{equation*}
$$

We introduce the identity that $\Psi \nabla^{2} P \equiv \nabla \cdot(\Psi \nabla P)-\nabla \Psi \cdot \nabla P$. Green's theorem converts the first term to a surface integral, so we now have

$$
\begin{equation*}
\iiint_{\mathcal{V}}\left(\nabla P \cdot \nabla \Psi-k^{2} P \Psi\right) d \mathcal{V}=-\iint_{\mathcal{S}}[(\bar{n} \cdot \nabla P) \Psi] d \mathcal{S} \tag{7.5.34}
\end{equation*}
$$

Note that the signs in this expression correspond to definition of $\bar{n}$ as the surface normal that points into the fluid domain.

The series in Eqs. (7.5.31) and (7.5.32) are used to describe $P$ and $\Psi$. To use the matrix form, the components of the gradient operator are placed in a column vector $\{\nabla\}$,

$$
\{\nabla\} \equiv\left[\begin{array}{ccc}
\frac{\partial}{\partial x_{n}} & \frac{\partial}{\partial y_{n}} & \frac{\partial}{\partial z_{n}} \tag{7.5.35}
\end{array}\right]^{\mathrm{T}}
$$

It will be noted that the gradient is expressed in terms of the local $x_{n} y_{n} z_{n}$ coordinate system, because that is the coordinate system used for the interpolating functions.

The pressure gradient is obtained by operating on Eq. (7.5.31), which leads to

$$
\{\nabla\} P=\left\{\begin{array}{l}
\partial P / \partial x_{n}  \tag{7.5.36}\\
\partial P / \partial y_{n} \\
\partial P / \partial z_{n}
\end{array}\right\}=\sum_{n=1}^{N}\{\nabla\}\left[\mathcal{N}^{(n)}\left(x_{n}, y_{n}, z_{n}\right)\right]\left[S^{(n)}\right]\{\hat{P}\}
$$

In this expression, only the interpolating functions depend on position. If $J$ is the total number of mesh points, then $\{\nabla\}\left[\mathcal{N}^{(n)}\left(x_{n}, y_{n}, z_{n}\right)\right]$ is a $3 \times J$ position-dependent array given by

$$
\{\nabla\}\left[\mathcal{N}^{(n)}\left(x_{n}, y_{n}, z_{n}\right)\right]=\left[\begin{array}{l}
\partial \mathcal{N}_{1}^{(n)} / \partial x_{n} \partial \mathcal{N}_{2}^{(n)} / \partial x_{n} \partial \mathcal{N}_{3}^{(n)} / \partial x_{n} \cdots  \tag{7.5.37}\\
\partial \mathcal{N}_{1}^{(n)} / \partial y_{n} \partial \mathcal{N}_{2}^{(n)} / \partial y_{n} \partial \mathcal{N}_{3}^{(n)} / \partial y_{n} \cdots \\
\partial \mathcal{N}_{1}^{(n)} / \partial z_{n} \partial \mathcal{N}_{2}^{(n)} / \partial z_{n} \partial \mathcal{N}_{3}^{(n)} / \partial z_{n} \cdots
\end{array}\right]
$$

Changing $P$ to $\Psi$ in the preceding gives the matrix representation of $\nabla \Psi$. A dot product is obtained as a product of row and column vectors, so the first term in Eq. (7.5.34) may be written as

$$
\begin{align*}
\nabla P \cdot \nabla \Psi= & (\{\nabla\} \Psi)^{\mathrm{T}}(\{\nabla\} P) \\
= & \sum_{m=1}^{N} \sum_{n=1}^{N}\{\eta\}^{\mathrm{T}}\left[S^{(m)}\right]^{\mathrm{T}}\left(\{\nabla\}\left[\mathcal{N}^{(m)}\left(x_{n}, y_{n}, z_{n}\right)\right]\right)^{\mathrm{T}}  \tag{7.5.38}\\
& \times\left(\{\nabla\}\left[\mathcal{N}^{(n)}\left(x_{n}, y_{n}, z_{n}\right)\right]\right)\left[S^{(n)}\right]\{\hat{P}\}
\end{align*}
$$

Both $\{P\}$ and $\{\eta\}$ are global parameters, so they may be brought outside the sum. Furthermore, the interpolating functions are zero outside their element. This means that all terms in the double sum for which the elements are different are identically zero. This reduces the double sum to a single sum over all elements, so that

$$
\begin{align*}
\nabla P \cdot \nabla \Psi= & \{\eta\}^{\mathrm{T}}\left[\sum_{n=1}^{N}\left[S^{(n)}\right]^{\mathrm{T}}\left(\{\nabla\}\left[\mathcal{N}^{(n)}\left(x_{n}, y_{n}, z_{n}\right)\right]\right)^{\mathrm{T}}\right.  \tag{7.5.39}\\
& \left.\times\left(\{\nabla\}\left[\mathcal{N}^{(n)}\left(x_{n}, y_{n}, z_{n}\right)\right]\right)\left[S^{(n)}\right]\right]\{\hat{P}\}
\end{align*}
$$

The scalar product $P \Psi$ is

$$
\begin{equation*}
P \Psi=\{\eta\}^{\mathrm{T}}\left[\sum_{n=1}^{N}\left[S^{(n)}\right]^{\mathrm{T}}\left[\mathcal{N}^{(n)}\left(x_{n}, y_{n}, z_{n}\right)\right]^{\mathrm{T}}\left[\mathcal{N}^{(n)}\left(x_{n}, y_{n}, z_{n}\right)\right]\left[S^{(n)}\right]\right]\{\hat{P}\} \tag{7.5.40}
\end{equation*}
$$

Now let us consider the right side of Eq. (7.5.34). This term is evaluated on the boundary of the acoustic domain, so that $\bar{n} \cdot \nabla P=-i \rho_{0} \omega V_{s}$. Thus this integrand may be written as

$$
\begin{equation*}
(\bar{n} \cdot \nabla P) \Psi=-i \omega \rho_{0}\{\eta\}^{\mathrm{T}} \sum_{n=1}^{N}\left[S^{(n)}\right]^{\mathrm{T}}\left[\mathcal{N}^{(n)}\left(x_{n}, y_{n}, z_{n}\right)\right]^{\mathrm{T}} V_{s}\left(x_{n}, y_{n}, z_{n}\right) \tag{7.5.41}
\end{equation*}
$$

The result of substituting these expressions into the Galerkin principle, Eq. (7.5.34), is

$$
\begin{equation*}
\{\eta\}^{\mathrm{T}}[A]\{\hat{P}\}-k^{2}\{\eta\}^{\mathrm{T}}[B]\{\hat{P}\}=i \omega \rho_{0}\{\eta\}^{\mathrm{T}}\{\hat{V}\} \tag{7.5.42}
\end{equation*}
$$

The matrices $[A]$ and $[B]$ are square and symmetric. Their definitions are

$$
\begin{align*}
& {[A]=\sum_{n=1}^{N} \iiint_{\mathcal{V}_{n}}\left[S^{(n)}\right]^{\mathrm{T}}\left(\{\nabla\}\left[\mathcal{N}^{(n)}\left(x_{n}, y_{n}, z_{n}\right)\right]\right)^{\mathrm{T}}\left(\{\nabla\}\left[\mathcal{N}^{(n)}\left(x_{n}, y_{n}, z_{n}\right)\right]\right)}  \tag{7.5.43}\\
& {[B]=\sum_{n=1}^{N} \iiint_{\mathcal{V}_{n}}\left[S^{(n)}\right] d \mathcal{V}}
\end{align*}
$$

The column vector $\{\hat{V}\}$ contains the effect of the surface velocity that generates the pressure field. It is given by

$$
\begin{equation*}
\{\hat{V}\}=\sum_{n=1}^{N} \iint_{\mathcal{S}}\left[S^{(n)}\right]^{\mathrm{T}}\left[\mathcal{N}^{(n)}\left(x_{n}, y_{n}, z_{n}\right)\right]^{\mathrm{T}} V_{s}\left(x_{n}, y_{n}, z_{n}\right) d \mathcal{S} \tag{7.5.44}
\end{equation*}
$$

The evaluation of $\{\hat{V}\}$ entails an integration over the surface, but most finite elements are on the interior. Those elements may be omitted from the sum. Also, it seldom is convenient to evaluate the integral for the actual $V_{s}$ function, especially in the context of a finite element computer code. Rather, the usual practice is to use a set of interpolating functions to represent the dependence of $V_{s}$ on the surface position in terms its value at those mesh points that lie on the surface.

Now we come to the crucial step. The test function $\Psi$ is arbitrary, which means that the $\eta_{j}$ coefficients are arbitrary. It is necessary that Eq. (7.5.42) be satisfied for any $\{\eta\}$, which will only be true if the coefficients of $\{\eta\}^{\mathrm{T}}$ in both sides of the equality are equal. The result is the standard form of the finite element equations of motion

$$
\begin{equation*}
\left[[A]-k^{2}[B]\right]\{\hat{P}\}=i \omega \rho_{0}\{\hat{V}\} \tag{7.5.45}
\end{equation*}
$$

Because $k^{2}$ is proportional to $\omega^{2}$, an individual who is conversant with the principles of structural dynamics might be tempted to call $[B]$ the inertia matrix, and $[A]$ the stiffness matrix. In fact, this is the opposite of the effect each represents. The kinetic energy density is $\left(\rho_{0} / 2\right) \bar{v} \cdot \bar{v}$, and $\bar{v}=-\operatorname{Re}\left(\left(1 / i \rho_{0} \omega\right) \nabla P\right)$. The integrand for $[A]$ contains the dot product of gradients, so it is the inertial effect. Similarly, the acoustic potential energy per unit volume is $p^{2} /\left(2 \rho_{0} c\right)$, and the integrand for $[B]$ contains a product of pressure functions. Hence, it is the stiffness term. The difference from the structural dynamics equations stems from the fact that we have used

Fig. 7.25 Finite element domain used to analyze radiation from the vibrating body at the center

pressure, which is a force-like quantity, as the basic variable, whereas displacement is the basic variable in structural dynamics.

The formulation leading to Eq.(7.5.45) is quite general, but it does not address the unique nature of a radiation problem in which energy is transported into an unbounded space. The problem is that it is not possible to divide an infinite domain into a finite number of elements. Individuals who have enormous computing power at their disposal have represented that domain as a very large sphere. This is the finite element model depicted in Fig. 7.25, where the radiating body is a hydrophone that consists of a cylinder with vibrating pistons at both ends. The mesh depicted there adheres strictly to a spherical coordinate grid, but it would likely be better to distort the mesh such that the radial lines in the vicinity of the hydrophone are close to being perpendicular to the surface.

Why is a spherical domain used? The answer lies in the farfield behavior of the field radiated by any finite-sized object. We know that in terms of spherical coordinates centered inside a radiating body, the pressure at a very large $r$ is a spherical wave with directivity. The complex pressure and particle velocity in that region are described by

$$
\begin{equation*}
P=\frac{f(\psi, \theta)}{r} e^{-i k r}, \quad V_{r}=\frac{P}{\rho_{0} c} \tag{7.5.46}
\end{equation*}
$$

The surface $\mathcal{S}$ is a composite of the exterior of the hydrophone $\mathcal{S}_{0}$ and the outer spherical surface $\mathcal{S}_{1}$. Correspondingly, Equation (7.5.44) for $\{\hat{V}\}$ is split into contributions from each surface. The surface velocity on the hydrophone is known. On $\mathcal{S}_{1}$, we use the farfield approximation. Recall that $V_{s}$ was defined to be positive if it is into the fluid, so $V_{s}=-V_{r}$ on the spherical boundary. Thus, the velocity coefficients
on $\mathcal{S}_{1}$ become

$$
\begin{align*}
\{\hat{V}\}_{1} & =\sum_{n=1}^{N} \iint_{\mathcal{S}_{1}}\left[S^{(n)}\right]^{\mathrm{T}}\left[\mathcal{N}^{(n)}\left(x_{n}, y_{n}, z_{n}\right)\right]^{\mathrm{T}} V_{s}\left(x_{n}, y_{n}, z_{n}\right) d \mathcal{S}  \tag{7.5.47}\\
& =\frac{1}{\rho_{0} c} \sum_{n=1}^{N} \iint_{\mathcal{S}_{1}}\left[S^{(n)}\right]^{\mathrm{T}}\left[\mathcal{N}^{(n)}\left(x_{n}, y_{n}, z_{n}\right)\right]^{\mathrm{T}} P\left(x_{n}, y_{n}, z_{n}\right) d \mathcal{S}
\end{align*}
$$

The pressure in the integrand is described by evaluating Eq. (7.5.31) on the spherical boundary. Doing so reduces the integrand to a product of interpolating functions, with $\{\hat{P}\}$ brought outside the integral. The result is an equation for $\{\hat{V}\}_{1}$ that depends on $\{\hat{P}\}_{1}$. Bringing these terms to the left side of Eq. (7.5.45) yields a set of equations for the mesh pressures in which the velocity coefficients on the radiating body are the inhomogeneous terms,

$$
\begin{equation*}
\left[[A]+k^{2}[B]+i k[C]\right]\{\hat{P}\}=i \omega \rho_{0}\{\hat{V}\}_{0} \tag{7.5.48}
\end{equation*}
$$

where the additional contribution to the coefficient matrix is

$$
\begin{equation*}
[C]=\sum_{n=1}^{N} \iint_{\mathcal{S}_{1}}\left[S^{(n)}\right]^{\mathrm{T}}\left[\mathcal{N}^{(n)}\left(x_{n}, y_{n}, z_{n}\right)\right]^{\mathrm{T}}\left[\mathcal{N}^{(n)}\left(x_{n}, y_{n}, z_{n}\right)\right]\left[S^{(n)}\right] d \mathcal{S} \tag{7.5.49}
\end{equation*}
$$

An important feature is that the term containing [C] is $90^{\circ}$ out-of-phase from the contributions of $[A]$ and $[B]$. As noted previously, $[A]$ and $[B]$ are associated, respectively, with the kinetic and potential energy of the fluid contained in the finite element model. Both effects are conservative. In contrast, $[C]$ represents a loss of energy, specifically, radiation damping resulting from power flowing out of the domain across the spherical boundary.

Although creating a model based on surrounding the radiating body with a very large sphere seems to be a reasonable procedure, it has a serious flaw. We do not know precisely where the farfield begins, but we do know that the sphere's radius $r_{1}$ must be sufficiently large that adjacent lines from a point on the outer surface to the vibrating body are essentially parallel. This means that the sphere's radius $r_{1}$ must be much greater than the largest dimension of the radiating body. The farfield approximation also requires that $k r_{1} \gg 1$, that is, $r_{1}$ must be much greater than an acoustic wavelength. Furthermore, the general requirement that the element size's be one-sixth of an acoustic wavelength, which was identified for boundary elements, also applies to finite elements. Meeting these criteria would require that the model contains many mesh points. That is why this approach typically requires substantial computational resources to analyze realistic radiation problems.

What should the rest of us do? Many concepts have been developed, all with the idea of shrinking the size of the sphere, and possibly, adjusting the shape of the outer boundary to conform more closely to the shape of the vibrating body. The
first modifies the spherical plane wave approximation to obtain a modified boundary condition. If the radius of the sphere is sufficiently large, setting $V_{r}=P /\left(\rho_{0} c\right)$ on the outer body will result in the reflection coefficient being zero. If a boundary had a zero reflection coefficient for any type of wave, we would say that is perfectly transmitting. The reflection coefficient for a plane wave obliquely incident on a locally reacting plane is

$$
\begin{equation*}
R=\frac{Z-Z_{1}}{Z+Z_{1}}, \quad Z_{1}=\frac{\rho_{0} c}{\cos \psi_{\mathrm{I}}} \tag{7.5.50}
\end{equation*}
$$

One could think of selecting $Z$ to minimize $|R|$ over a range of $\psi_{\mathrm{I}}$, and then using that $Z$ as the boundary condition for the exterior surface $\mathcal{S}$ of the finite element mesh. That surface need not be spherical, so the approximate transmitting boundary condition would be

$$
\begin{equation*}
P=-Z \bar{n} \cdot \bar{V} \tag{7.5.51}
\end{equation*}
$$

where the negative sign results from defining $\bar{n}$ to point into the domain represented by the mesh.

This approximation does not work too well, so the next idea is to introduce dissipation, in which case the result will be an absorbing boundary condition. This can be done by allowing $Z$ to be complex, or by introducing a dissipation operator into the plane wave approximation. A simple absorbing boundary condition would be

$$
\begin{equation*}
\rho_{0} c \bar{n} \cdot \bar{V}=-(P+\alpha \bar{n} \cdot \nabla P) \tag{7.5.52}
\end{equation*}
$$

where the value of $\alpha$ would be selected to minimize the signal that is reflected back into the acoustic domain. The absorbing boundary conditions that are used are much more complicated than these. An archival work is the paper by Bayliss and Turkel ${ }^{24}$; a more accessible description is the paper by Assaad et al. ${ }^{25}$

The latest development is the concept of an infinite element. Such an element uses the basic form of a radially diverging wave as the basis for its shape functions. These elements are placed outside the fluid domain that is described by regular finite elements. Most derivations have considered this domain to be spherical, which is the configuration depicted in Fig. 7.26. The exception is the work of Burnett ${ }^{26}$ in which the outer boundary of the fluid domain is a prolate spheroid, which can enclose a long slender object without extending far in the broadside direction.

As shown in Fig. 7.26, some mesh points lie on the outer boundary $\mathcal{S}_{1}$ of the finite element mesh. The field within the element is taken to be an outgoing spherical wave with directivity. This dependence is expressed in terms of spherical coordinates centered on the sphere, which leads to

[^19]Fig. 7.26 Configuration of an infinite element used to approximate the free field in a finite element model

$$
\begin{align*}
& P=\frac{e^{-i k r}}{r} \sum_{m=0}^{M} \frac{F_{m}(\psi, \theta)}{r^{m}}
\end{align*}
$$

Several methods have been used to set the $F_{m}(\psi, \theta)$ functions. One, described by Burnett, ${ }^{27}$ selects the $m=0$ function to match the regular finite element functions in the two transverse directions, and then finds the others by fitting the expression to the angular terms in the Laplacian operator. A different derivation based on a mapping of the cylindrical coordinates into a hyperbolic grid was used by Astley et al. ${ }^{28}$ In any event matching, the standard interpolation form in Eq. (7.5.31) to the diverging spherical wave form leads to the interpolation functions for the infinite element. At that juncture, the contribution of the infinite elements to $[A]$ and $[B]$ is folded into those of the regular finite elements.

The advantages of the finite element formulation might not be apparent at this juncture, especially because the one-sixth wavelength requirement also applies to a finite element model. A boundary element model only entails a mesh on the surface, whereas finite elements discretize the three-dimensional domain of the fluid. Hence, a finite element model will have more mesh points, and therefore more equations to solve, than a comparable boundary element model. The advantageous aspects of the finite element model follow from its basic features. First, the coefficient matrices [ $A$ ], $[B]$, and $[C]$ are constants independent of frequency. Thus, in the likely event that it is necessary to perform a frequency sweep, the basic terms for the equations of motion are only evaluated once. In contrast, the coefficient matrices for a boundary element model depend on frequency. Another difference is that the integrals for the evaluation of coefficient matrices can be derived as standard formulas for a specified set of interpolation functions. In contrast, the integrals for boundary element coefficients contain singularities and therefore constitute a more formidable task to program and evaluate. Some individuals find the absence of the "forbidden frequency" issue to be a major asset of finite elements.

[^20]Perhaps the most significant difference between finite elements and boundary elements may be found in the fundamental mathematical properties of their equations. As was noted earlier, the coefficient matrix for boundary elements is full. In contrast, the mesh points for a finite element model are shared by only a few elements, so only a few elements are directly coupled. This means that the coefficient matrices for a finite element model will be banded around the diagonal. Many highly efficient methods have been developed to solve equations whose system matrices are banded. Thus, although the finite element model has more mesh points, it tends to require less computational resources.

An interesting feature of the finite element technique is that all developments, including absorbing boundary conditions and infinite elements, have analogous formulations in the time domain. Whether this is useful depends on what properties of the response one wishes to obtain. Often, a complex frequency response is the desired result. Extracting time domain data from a frequency domain simulation would require FFT processing, thereby adding to the computational effort. On the other hand, solving time domain equations of motion must be done with a numerical differential equation solver, which require a great deal of computational resources if it is necessary to a large number of coupled equations. In any event, the time domain formulation lies more in the research arena because some features are still being explored.

It is worth noting that the finite element formulation is equally valid for the analysis of a cavity. The same is true for the boundary element formulation, if the direction of the surface normal is reversed. Several commercial codes of both types are available for such applications. Both are simplified by the fact that a cavity is a finite domain. Hence, there is no need for absorbing boundary conditions or infinite elements, and the forbidden frequency issue is irrelevant.

### 7.6 Homework Exercises

Exercise 7.1 The farfield radiation of a vibrating sphere is observed to match the pattern of a longitudinal quadrupole. The fluid is air and the radius of the sphere is 100 mm . When the frequency is 1.2 kHz , the maximum pressure at 5 m from the center of the sphere is 0.3 Pa . What vibration of the sphere's surface would produce a pressure field having these attributes?

Exercise 7.2 The surface velocity on a sphere is given by $v_{r}=\operatorname{Re}\left[V_{4} \cos (4 \phi)\right.$ $\exp (-i \omega t)]$. Derive an expression for the complex pressure amplitude in the radiated field. Evaluate $\left|P / \rho_{0} c V_{4}\right|$ along the polar axis, $\phi=0$. Plot the result as a function of $k r$ when $k a=4$ and $k a=24$. Based on these graphs, identify the value of $k r$ at which the farfield approximation is valid for each $k a$.

Exercise 7.3 A sphere vibrates over a band surrounding its equator, specifically, $v_{r}=\operatorname{Re}\left[V_{0} \exp (i \omega t)\right]$ if $\pi / 3<\psi<2 \pi / 3, v_{r}=0$ otherwise. Derive an expression
for the acoustic pressure in the farfield. From this, evaluate the directivity at $k a=1$, 10, and 50. Interpret these results relative to what one would expect qualitatively.

Exercise 7.4 Scanning the farfield of a vibrating sphere has led to identification of the coefficients of a spherical harmonic series describing the polar angle dependence of the pressure at a specific radial distance $r_{1}=k a^{2}$. Specifically, it is known that only spherical harmonics above $m=8$ are insignificant, with the contribution of the lower harmonics being

$$
P\left(r_{1}, \psi\right)=P=\rho_{0} c^{2} \sum_{m=0}^{8} C_{m} P_{m}(\cos \psi), \quad C_{m}=\frac{1}{(m+0.5)^{2}}
$$

Determine the radial velocity on the surface of the sphere for $k a=2$ and $k a=20$. Compare the polar pattern of the surface velocity to that of the measured pressure field.

Exercise 7.5 Due to measurement error, the coefficients $C_{m}$ in Exercise 7.4 have a $\pm 20 \%$ error in amplitude and a $\pm 40^{\circ}$ error in phase relative to the nominal value given there. The errors for each coefficient have a uniform probability distribution. Create a new set of coefficients by adding the random error to each of the stated values. Determine the radial velocity for $k a=0.5,4$, and 20 at the surface corresponding contaminated coefficients. Compare that result to the surface velocity when the coefficients have their nominal values. What conclusions can be drawn regarding this inverse identification process?

Exercise 7.6 A vibrating sphere of radius $a$ is surrounded by a polymeric material that has been molded into a sphere of radius $b$ that is concentric with the inner sphere. Outside these spheres, the fluid is water. The shear strength of the polymeric material is sufficiently small that the material may be considered to be a liquid whose density is $\rho_{1}$ and whose sound speed is $c_{1}$. The vibration of the inner sphere is radially symmetric, so that $v_{r}=v_{a} \sin (\omega t)$ at $r=a$. Derive an expression for the pressure in the water. Hint: The pressure and radial particle velocity must be continuous at $r=b$.

Exercise 7.7 A hemispherical balloon is fastened on its edges to a rigid baffle. A jet of air internally induces a vibration that has a bulbous distribution given by $v_{r}=v_{0} \cos (2 \psi) \sin (\omega t)$ for $r=a$ and $0 \leq \psi \leq \pi / 2$. (a) Derive a spherical harmonic series describing the radiated pressure field. (b) Evaluate the farfield directivity for $k a=0.5,5$, and 20 .

Exercise 7.8 A hemispherical balloon is fastened on its edges to an infinite pressurerelease baffle. A jet of air internally induces a vibration that has a bulbous distribution given by $v_{r}=v_{0} \cos (2 \psi) \sin (\omega t)$ for $r=a$ and $0 \leq \psi \leq \pi / 2$. (a) Derive a spherical harmonic series describing the radiated pressure field. (b) Evaluate the farfield directivity for $k a=0.5,5$, and 20 .

Exercise 7.9 A hemispherical steel shell is fastened at its rim, $\psi=90^{\circ}$, to a rigid baffle. The attachment is such that displacement parallel to the baffle is possible, but displacement perpendicular to the baffle is not. In terms of the shell displacements, these conditions require that $u=0$ at $\psi=\pi / 2$. The internal pressure is a uniform distribution that fluctuates harmonically, so that $p_{s}(\psi, t)=P_{0} \cos (\omega t)$. The shell's diameter is a factor of 120 greater than its thickness, its density is $5 \rho_{0}$, and the sound speed of the shell is $c_{e}=3 c$. (a) Derive spherical harmonic series describing the radiated pressure field and the shell displacement. (b) Evaluate the farfield directivity for $k a=0.5,5$, and 20 .

Exercise 7.10 A hemispherical steel shell is fastened at its rim, $\psi=90^{\circ}$, to a pressure-release baffle. The attachment is such that displacement perpendicular to the baffle is possible, but radial displacement parallel to the baffle is not. In terms of the shell displacements, these conditions require that $w=0$ at $\psi=\pi / 2$. The internal pressure is a uniform distribution that fluctuates harmonically, so that $p_{s}(\psi, t)=P_{0} \cos (\omega t)$. The shell's diameter is a factor of 120 greater than its thickness, its density is $5 \rho_{0}$, and the sound speed of the shell is $c_{e}=3 c$. (a) Derive spherical harmonic series describing the radiated pressure field and the shell displacement. (b) Evaluate the farfield directivity for $k a=0.5,5$, and 20.

Exercise 7.11 It is suspected that a long cylindrical section of an HVAC duct is vibrating excessively, thereby radiating an unacceptable sound level. The diameter of this section is 240 mm . Measurements at a distance of 2.5 m from the centerline indicate that the sound pressure level is $95 \mathrm{~dB} / / 20 \mu \mathrm{~Pa}$, and that the wavenumber is $\bar{k}^{\prime}=20 \bar{e}_{R}+1.6 \bar{e}_{\theta}+15 \bar{e}_{z} \mathrm{~m}^{-1}$. What properties of the duct's vibration can be deduced from these measurements?

Exercise 7.12 An extremely long pipeline is submerged in a deep region of the ocean. The diameter of the pipe is 700 mm . A disturbance at a distant end generates a flexural wave in the pipe at wavenumber. A reference location for the radiated pressure is $R=600 \mathrm{~mm}, z=0$. The pressure measured at this location is $940 \cos (3770 t+$ 0.1160 ). Then, the position of the hydrophone is shifted gradually. Measurements at various circumferential angles lead to the conclusion that field is axisymmetric. The radial distance $R$ is increased gradually. It is found that $R_{2}=2.48895 \mathrm{~m}$ is the smallest distance greater than $R_{1}=0.6 \mathrm{~m}$ at which the pressure at the two locations is $180^{\circ}$ out-of-phase. After this measurement is made, the hydrophone is shifted axially from the reference location, with $R$ held constant. This new measurement indicates that the pressure at $R_{3}=0.6, z_{3}=1.6017$ also is $180^{\circ}$ out-of-phase from the reference signal. (a) Determine the axial and radial wavenumbers. (b) Determine the complex amplitude of the surface vibration.

Exercise 7.13 A pressure field is the superposition of a plane wave that propagates in the axial direction of a cylindrical pipe and an axisymmetric cylindrical standing wave. Specifically, the pressure at an arbitrary field point is $p=$ $B H_{0}(\kappa R) \cos (\pi z / L) \cos (\omega t)+C \cos (\omega t-k z)$. Derive an expression for the timeaveraged power per unit axial length radiated by the cylinder.

Exercise 7.14 The surface normal velocity on a very long cylinder consists of a bulging motion on one side of the circumference, with a sinusoidal variation in the axial direction. The specific form is $\bar{v} \cdot \bar{e}_{R}=V_{0} \cos (\theta) \cos (\pi z / L)$ if $|\theta| \leq \pi / 2, \bar{v}$. $\bar{e}_{R}=0$ if $\pi / 2<|\theta| \leq \pi$. The wavelength is twice the cross section's radius, $L=2 a$, and the frequency is $k a=0.75 \pi$. (a) Derive an expression for the radiated pressure as a function of $R, z$, and $\theta$. (b) Evaluate the directivity function $(R / a)^{1 / 2}\left|P_{\mathrm{ff}}\right| /\left(\rho_{0} c V_{0}\right)$ as a function of $\theta$ in the plane $z=0$. (c) Evaluate $(R / a)^{1 / 2}|P| /\left(\rho_{0} c V_{0}\right)$ as a function of $R$ along the line defined by $\theta=0, z=0$. From that result, estimate where the farfield begins.

Exercise 7.15 A cylindrical pipeline is supported at intervals spaced at distance $L$. It transports a liquid whose flow velocity is high, with the consequence that the cross sections undergo a translational displacement in a fixed transverse direction, which is defined to be $x$. In addition, pressure fluctuations within the flowing liquid induce a uniform expansion and contraction of a cross section. The result is that the normal displacement on the cylinder's surface is $w=\left[W_{0}+W_{1} \cos (\pi z / L) \cos \theta\right] \sin (\omega t)$, where $z$ is the axial distance from a reference point. The cylinder's diameter is $2 a$ and the half-wavelength is $L=a$. Derive an expression for the dependence of the pressure on the radial distance $R$, axial distance $z$, and circumferential angle $\theta$ measured from the $x$-axis. It may be assumed for this analysis that end effects are negligible. (b) Postprocessing of data acquired in an experiment indicates that the pressure amplitude when $W_{0}$ and $W_{1}$ are nonzero is the twice the amplitude when $W_{0}=0$. The location for this observation is situated on the transverse line for which $\theta=0$ and $z=0$. Determine the ratio $W_{0} / W_{1}$ for each of the following combinations of frequency and location: (1) $k a=1$ and $R=a$, (2) $k a=1$ and $R=10 a$, (3) $k a=10$ and $R=a$, (4) $k a=10$ and $R=10 a$.

Exercise 7.16 The surface of a very long cylinder whose radius is $a$ consists of a sequence of circular bands of length $L$. These bands are piezoceramic elements, each of which vibrates axisymmetrically at frequency $\omega$ with no axial variation. Each band's velocity is $180^{\circ}$ out-of-phase from those to which it is adjacent. The result is that the surface velocity has a square wave pattern in the axial direction, given by $v_{R}=\operatorname{Re}(V(z) \exp (i \omega t))$, where the complex amplitude is $V(z)=V_{0}$ if $0<z<L, V(z)=-V_{0}$ if $L<z<2 L, V(z)=V(z \pm 2 L)$. Derive an expression for the farfield pressure. Plot $(R / a)^{1 / 2}\left|p /\left(\rho_{0} c V_{0}\right)\right|$ as a function of $z$ for the case where $k a=1$ and $k a=10$. Does either result resemble the square wave pattern on the surface?

Exercise 7.17 The field radiated by a sphere that executes an oscillatory rigid body translation has a known solution as a dipole, see Sect.6.5.2. Use the results derived there to evaluate the surface pressure when a sphere's surface velocity is $\operatorname{Re}\left(V_{1} \exp (i \omega t)\right) \cos \psi$. Then, use this surface response to formulate the KHIT for the pressure at an arbitrary field point. In the case where $k a=5$, evaluate the result of that formulation along the polar axis, $\psi=0$, for $0<r<4 a$. Compare the result to the analytical description, Eq. (6.5.102).

Exercise 7.18 A cylindrical rod whose length is $L$ and diameter is $2 a$ vibrates like a simply supported beam. This means that all points on a cross section displace by the same amount in a fixed transverse direction. In terms of cylindrical coordinates whose origin is at the midpoint, the resulting velocity of a surface point is $\bar{v}=\operatorname{Re}\left(V_{1} \exp (i \omega t)\right) \sin (\pi z / L)\left(\bar{e}_{r} \cos \theta+\bar{e}_{\theta} \sin \theta\right)$. It is desired to ascertain the farfield directivity, so it is not acceptable to approximate the rod as an infinite length cylinder. Use the formulation in Example 7.10, in which the radiation impedance for an infinite cylinder is used to estimate the surface pressure corresponding to the stated surface velocity. An integral expression for the far field pressure can be obtained by substituting the surface velocity and approximate surface pressure into the KHIT. (a) Derive this integral expression. (b) Use analytical or numerical methods to evaluate and plot $|R| P_{\mathrm{ff}} /\left(\rho_{0} c V_{1}\right)$ in the plane $\theta=0$ as a function of the polar angle $\psi$. Parameters are $k a=1$ and $10, L / a=6$.

Exercise 7.19 The rigid disk in the sketch executes an oscillatory transverse vibration as a rigid body. The normal velocity on the front surface, whose normal is $\bar{e}_{z}$, is $\bar{v}_{\text {front }} \cdot \bar{n}=\operatorname{Re}\left(V_{1} \exp (i \omega t)\right)$, whereas on the back surface, it is $\bar{v}_{\text {back }} \cdot \bar{n}=$ $\operatorname{Re}\left(-V_{1} \exp (i \omega t)\right)$. It is very thin, so the surface response on the edge has negligible effect. Because the front and back motions are $180^{\circ}$ out-of-phase, but otherwise alike, the surface pressures on opposite sides also must be $180^{\circ}$ out-of-phase. Furthermore, the pressure distributions must be axisymmetric relative to the $z$-axis, so it must be that $p_{\text {front }}=-p_{\text {back }}=\operatorname{Re}(P(R) \exp (i \omega t))$. Another consideration comes from the theory of diffraction. According to it, the zero thickness approximation leads to a discontinuity at $R=a$. This requires that the pressure at the edge be zero with an infinite slope, $d p_{\text {front }} / d R=0$. A trial function fitting these requirements is $P_{\text {front }}=P_{\mathrm{C}}\left(1-R^{2} / a^{2}\right)^{1 / 2}$. (a) Show that using this guess to formulate the farfield version of the KHIT, Eq. (7.4.18), leads to

$$
P_{\mathrm{ff}}\left(\bar{x}_{0}\right)=i k \frac{e^{-i k r}}{r}(\cos \psi) \int_{0}^{a} P_{\mathrm{C}}\left(1-\frac{R^{2}}{a^{2}}\right)^{1 / 2} J_{0}(k R \sin \psi) R d R
$$

(b) In the limit as $k a \rightarrow 0$, this expression should yield the same result as Eqs. (7.4.38) and (7.4.39), which is the (low-frequency) multipole expansion for a transversely oscillating disk. Use this fact to derive an expression for $P_{\mathrm{C}}$.


Exercise 7.19

Exercise 7.20 The sketch shows a rigid box mounted on a vertical shaft. Its rotation rate is $\Omega=\Omega_{0} \sin (\omega t)$. Identify the restrictions that must be imposed on $\Omega_{0}$ and $\omega$ in order that a linearized analysis be valid. Given that these restrictions are met, describe the multipole expansion that would fit the field radiated by the box. In what directions can the pressure amplitude be expected to have a maximum value? In what directions can the pressure be expected to be zero? Estimate the largest pressure amplitude that would be observed at a fixed distance $r$ from the origin.


Exercise 7.20
Exercise 7.21 The surface Helmholtz integral equation was derived in Sect. 7.5.2 for the case of radiation by letting $\bar{x}_{0}$ be an exterior field point that approaches the surface $\mathcal{S}$. Show that the same result is obtained if $\bar{x}_{0}$ is a field point inside the radiating body.

Exercise 7.22 The field radiated by a sphere that executes an oscillatory rigid body translation has a known solution as a dipole, see Sect. 6.5.2. Thus, it offers the opportunity to test an implementation of the source superposition formulation. The arrangement of sources and field points is described in the sketch. There are four sources that are situated along the $z$-axis, which is the direction of the translational velocity $\bar{v}=V_{1} \sin (\omega t) \bar{e}_{z}$. These points are spaced at $0.4 a$ centered on the sphere's center. There are nine surface points spaced equally along a meridian. (a) Explain why it is preferable that all of sources be situated on the $z$-axis. (b) Explain why it would be incorrect to use source or surface points at the same polar angle along more than one meridian. (c) Without solving the superposition equations, anticipate how the source strengths left of the center should be related to those to the right of the center. (d) Formulate and solve the source superposition equations for the case where $k a=8$. Compare the pressure on the surface at $\psi=30^{\circ}$ to the dipole solution.


Exercise 7.22

Exercise 7.23 The sketch shows a model for an underwater projector. It is a cylinder whose surfaces and one end are stationary. The end that is visible in the sketch executes a translational oscillation, such that all points on that surface vibrate in the axial direction according to $v_{z}=V \cos (\omega t)$. The task is to use the source superposition method to evaluate the radiation. The eventual objective would be to evaluate the radiated pressure field, but here it is sufficient to evaluate the source strengths. The arrangement of sources and surface points is such that both are spaced at onefifth of the cylinder's length centered on the midpoint in the axial direction, with an additional six field points, three on each end. (a) Explain why it is preferable that all sources be situated on the $z$-axis. (b) Explain why it would be incorrect to use source points at the axial location along other circumferential angles. (c) Explain why it would be unwise to locate a surface point at the junction of the cylinder and its end. (d) For the case where $L=3 a$ and $k a=4$, determine the source strengths as nondimensional quantities, $\hat{Q}_{n} /\left(a^{2} V\right)$.


Exercise 7.23

## Chapter 8 <br> Radiation from a Source in a Baffle

Section 3.5.1 introduced a correction for the pressure-release condition at the open end of a one-dimensional waveguide. The basis for that development is the radiation properties of a circular piston flush mounted in a wall whose extent is large. The wall is referred to as a baffle. A piston serves well as a model of real transducers that might be piezoceramics in an underwater projector. It may be employed as an approximate representation of the movable cone of a loudspeaker. The term "piston" is intended to convey the notion that all points in the surface have the same velocity. Our study will concentrate on such situations, but allowance will be made in the development for more general velocity distributions.

Because the concern here is with the signal that is generated by a type of surface vibration, it could logically be included in the previous chapter. However, the phenomena that arise are extremely relevant to a number of every day systems. Consequently, many aspects of radiation from a transducer in a planar baffle have been thoroughly analyzed. The subject appears in its own chapter as a way of emphasizing its importance.

The KHIT is the fundamental principle that is the foundation for the developments that follow. Rayleigh identified a Green's function that removes the need to know the surface pressure, with the result that the pressure at a field point is described as a surface integral that requires specification of the normal velocity. Given that the KHIT we derived governs the frequency domain, that is the description with which we begin, but we will also examine a time-domain solution.

### 8.1 The Rayleigh Integral

Figure 8.1 depicts the system to be analyzed. We wish to determine the pressure $\operatorname{Re}\left(P\left(\bar{x}_{0}\right) e^{i \omega t}\right)$ at an arbitrary field point $\bar{x}_{0}$ that results from a specified normal velocity distribution $\operatorname{Re}\left(\bar{n}\left(\bar{x}_{s}\right) \cdot \bar{V}\left(\bar{x}_{s}\right) e^{i \omega t}\right)$ on $\mathcal{S}$. This plane is the boundary of a

Fig. 8.1 An aribitrary velocity distribution over points $\bar{x}_{s}$ on a flat plane that is the boundary $\mathcal{S}$ of a half-space

half space. The origin of the $x y z$ coordinate system is situated on $\mathcal{S}$, with the $z$-axis pointing into the fluid domain. The only requirement imposed on $\bar{V}\left(\bar{x}_{s}\right)$ is that it approach zero with increasing distance in all directions from the origin.

The KHIT entails integration over the surface that bounds the domain $\mathcal{V}$ in which the signal occurs. In the present situation that domain is a half space. The plane is only a part of the boundary, with the other part extending to infinity. The concept of an enclosing virtual surface allows us to apply the KHIT to this configuration. In Sect.7.4.2, a sphere that surrounds the vibrating body was used to extend the KHIT to radiation problems. Here, we use a hemisphere $\mathcal{S}_{0}$ centered on the origin, whose radius $r_{0}$ is allowed to grow without limit. The reasoning that previously led to recognition that the contribution of $\mathcal{S}_{0}$ to the KHIT is negligible if $r_{0}$ is very large also applies here. The consequence is that the KHIT only requires integration over the planar surface $\mathcal{S}$.

The discussion in Sect. 6.4.2 showed that in addition to the free-space Green's function, other functions that satisfy boundary conditions can be defined. If we can find a Green's function for which $\bar{n}\left(\bar{x}_{s}\right) \cdot \nabla G\left(\bar{x}_{0}, \bar{x}_{s}\right)$ is identically zero, then the integrand in the KHIT will not contain the surface pressure. Such an arrangement would make it possible to determine the pressure field from a direct evaluation of the integral.

The first step in identifying the requisite Green's function is to remove any ambiguity in what we seek. The position of the field point $\bar{x}_{0}$ is represented by its coordinates $x_{0}, y_{0}$, and $z_{0}$ and a source point $\bar{x}$, not necessarily on the surface, is situated at coordinates $x, y$, and $z$. If we reduce $z$ to zero, the source point will be the surface point $\bar{x}_{s}$. In other words, $\bar{x}=\bar{x}_{s}+z \bar{e}_{z}$. The gradient in the KHIT is evaluated at the source point, and we want the source point to be on the surface. Thus, we seek a Green's function that satisfies

$$
\begin{equation*}
\bar{n}\left(\bar{x}_{s}\right) \cdot \nabla_{s} G\left(\bar{x}_{0}, \bar{x}_{s}\right) \equiv \lim _{z \rightarrow 0} \frac{\partial}{\partial z} G\left(\bar{x}_{0}, \bar{x}\right)=0 \tag{8.1.1}
\end{equation*}
$$

Our experience with a point source in a half space suggests where to search for the desired function. According to the method of images, the field associated with a point source situated above a rigid plane is obtained by adding an equal point source at the same distance below the plane. Figure 8.2 depicts a point source at distance $z$ above the surface and its image at distance $z$ below the surface. The projection of $\bar{x}^{\prime}$ onto the surface is also $\bar{x}_{s}$.

Fig. 8.2 Green's function for a half space bounded by a rigid planar boundary. The source is at $\bar{x}$ and the image is at $\bar{x}^{\prime}$


The mass acceleration of the point source is defined to be one. The image location is $\bar{x}^{\prime}=\bar{x}_{s}-z \bar{e}_{z}$. The sum of this source and its image is a Green's function for the half space. This is so because the field radiated by the pair satisfies the inhomogeneous Helmholtz equation, $\nabla^{2} G+k^{2} G=-\delta\left(\bar{x}_{0}-\bar{x}\right)$ for any $\bar{x}$ and $\bar{x}_{0}$ within the half space, as well as the radiation condition, and the boundary condition of zero normal velocity on the surface. To distinguish the half-space function $G$ from the free-space Green's function let us use a subscript " fs " to identify the latter. Thus, the function we have constructed is

$$
\begin{equation*}
G\left(\bar{x}, \bar{x}_{0}\right)=G_{\mathrm{fs}}\left(\bar{x}_{s}+z \bar{e}_{z}, \bar{x}_{0}\right)+G_{\mathrm{fs}}\left(\bar{x}_{s}-z \bar{e}_{z}, \bar{x}_{0}\right) \tag{8.1.2}
\end{equation*}
$$

We wish to reduce $z$ to zero. Before we do so, let us evaluate $\partial G / \partial z$ when $z$ is not zero. The radial distances to the field point from the source and image are

$$
\begin{gather*}
r=\left|\bar{x}_{s}+z \bar{e}_{z}-\bar{x}_{0}\right|=\left[\left(x-x_{0}\right)^{2}+\left(y-y_{0}\right)^{2}+\left(z-z_{0}\right)^{2}\right]^{1 / 2}  \tag{8.1.3}\\
r^{\prime}=\left|\bar{x}_{s}-z \bar{e}_{z}-\bar{x}_{0}\right|=\left[\left(x-x_{0}\right)^{2}+\left(y-y_{0}\right)^{2}+\left(-z-z_{0}\right)^{2}\right]^{1 / 2}
\end{gather*}
$$

Because the free-space Green's function depends only on the radial distance between the pair of points, the $z$ derivative may be evaluated according to

$$
\begin{align*}
\frac{\partial}{\partial z} G\left(\bar{x}_{0}, \bar{x}\right) & =\frac{\partial r}{\partial z} \frac{d}{d r} G_{\mathrm{fs}}\left(\bar{x}_{s}+z \bar{e}_{z}, \bar{x}_{0}\right)+\frac{\partial r^{\prime}}{\partial z} \frac{d}{d r^{\prime}} G_{\mathrm{fs}}\left(\bar{x}_{s}-z \bar{e}_{z}, \bar{x}_{0}\right)  \tag{8.1.4}\\
& =\frac{z-z_{0}}{r} \frac{d}{d r} G_{\mathrm{fs}}\left(\bar{x}_{s}+z \bar{e}_{z}, \bar{x}_{0}\right)+\frac{z+z_{0}}{r^{\prime}} \frac{d}{d r^{\prime}} G_{\mathrm{fs}}\left(\bar{x}_{s}-z \bar{e}_{z}, \bar{x}_{0}\right)
\end{align*}
$$

When we reduce $z$ to zero, both $\bar{x}$ and $\bar{x}^{\prime}$ become $\bar{x}_{s}$. Consequently, $G_{\mathrm{fs}}\left(\bar{x}_{s}+z \bar{e}_{z}, \bar{x}_{0}\right)$ and $G_{\mathrm{fs}}\left(\bar{x}_{s}-z \bar{e}_{z}, \bar{x}_{0}\right)$ both become $G_{\mathrm{fs}}\left(\bar{x}_{s}, \bar{x}_{0}\right)$, so that $d G_{\mathrm{fs}} / d r^{\prime}$
approaches $d G_{\mathrm{fs}} / d r$. Furthermore, in the limit, the fractions preceding each $G_{\mathrm{fs}}$ term in the above expression become $-z_{0} / r$ and $z_{0} / r$. Thus, we conclude that the Green's function defined in Eq.(8.1.2) with $z=0$ satisfies Eq.(8.1.1). In other words, the Green's function we seek is

$$
\begin{equation*}
G\left(\bar{x}_{0}, \bar{x}_{s}\right)=2 G_{\mathrm{fs}}\left(\bar{x}_{0}, \bar{x}_{s}\right) \tag{8.1.5}
\end{equation*}
$$

The result of using this Green's function to form KHIT is

$$
\begin{equation*}
P\left(\bar{x}_{0}\right)=2 i \omega \rho_{0} \iint_{\mathcal{S}} \bar{n}\left(\bar{x}_{s}\right) \cdot V\left(\bar{x}_{s}\right) G_{\mathrm{fs}}\left(\bar{x}_{0}, \bar{x}_{s}\right) d \mathcal{S} \tag{8.1.6}
\end{equation*}
$$

In retrospect, Eq. (8.1.5) could have been identified heuristically. A free-space Green's function represents the limit as its radius shrinks to zero of a radially symmetric spherical source, whose mass acceleration is unity. When this sphere is placed on a rigid surface at $\bar{x}_{s}$, its mass flow is confined to the hemisphere above the boundary. If the mass flow is held at a unit value, the particle velocity on the hemisphere is double that for the full sphere in free space. This results in doubling the radiated pressure relative to what would have been obtained if the surface were not present. In regard to the gradient property, consider a spherical source in free space at location $\bar{x}_{s}$. Changing the $z$ coordinate of the source will change the distance to a field point, and therefore the pressure at that point. In contrast, if a hemispherical source is situated on the boundary, moving it to a small distance above the boundary moves its image below the boundary by the same distance. The change in the distance from the image to the field point is the negative of the change in the distance from the source to the field point. The pressure changes for each cancel.

Equation (8.1.6) is valid for two-dimensional, as well as three-dimensional models. This is so because the only place where dimensionality was considered was in the representation of $r_{1}$ and $r_{2}$, and setting $y_{s}$ and $y_{0}$ to zero for a twodimensional geometry would not alter any result. However, our primary interest is in three-dimensional configurations. When $G_{\mathrm{fs}}\left(\bar{x}_{s}, \bar{x}_{0}\right)$ is specified as the threedimensional version, the form of Eq.(8.1.6) that results is known as the Rayleigh integral. ${ }^{1}$ It is

$$
\begin{equation*}
P\left(\bar{x}_{0}\right)=\frac{i \omega \rho_{0}}{2 \pi} \iint_{\mathcal{S}} V_{s}\left(\bar{x}_{s}\right) \frac{e^{-i k\left|\bar{x}_{0}-\bar{x}_{s}\right|}}{\left|\bar{x}_{0}-\bar{x}_{s}\right|} d \mathcal{S} \tag{8.1.7}
\end{equation*}
$$

where $V_{s}\left(\bar{x}_{s}\right)$ is the velocity normal to a surface at location $\bar{x}_{s}$,

$$
\begin{equation*}
V_{s}\left(\bar{x}_{s}\right)=\bar{n}\left(\bar{x}_{s}\right) \cdot \bar{V}\left(\bar{x}_{s}\right) \tag{8.1.8}
\end{equation*}
$$

[^21]The Rayleigh integral describes situations where $\mathcal{S}$ is the planar boundary of an infinite half space. We will see in Sect. 8.4 that it may be used for a finite baffle, provided that the baffle extends at least one acoustic wavelength beyond the portion of $\mathcal{S}$ that moves. However, it has been used as an approximation when $\mathcal{S}$ is not planar. Doing so requires a thorough verification through computational modeling or experiment because it is not justified by the derivation. This issue will be examined in Example 8.4.

A detailed mapping of the pressure field requires that the Rayleigh integral be evaluated numerically. How to perform such an evaluation is addressed in Sect. 8.4. However, our initial explorations will examine analytical solutions that are descriptive of the field in specific regions.

### 8.2 Farfield Directivity

Suppose the surface vibration is confined to a finite portion of $\mathcal{S}$, which will be designated as $\mathcal{S}_{v}$. If the distance from the origin to the field point $\bar{x}_{0}$ is much greater than the maximum distance between two points in $\mathcal{S}_{v}$, then we may invoke the farfield approximation of the radial distance from a surface point to $\bar{x}_{0}$.

### 8.2.1 Cartesian Coordinate Description

We consider first the case where the outline of the vibrating region is best described in terms of the Cartesian coordinates $x$ and $y$ that lie on the surface. A point in $\mathcal{S}_{v}$ is $\bar{x}_{s}$, and $\bar{x}_{0}$ is a field point. The farfield approximation requires that $\left|\bar{x}_{0}\right|$ be much greater than the largest $\left|\bar{x}_{s}\right|$. When this condition applies, the radial line from $\bar{x}_{0}$ to any $\bar{x}_{s}$ is essentially parallel to the radial line from $\bar{x}_{0}$ to the origin of $x y z$, so that the length of radial lines differs by the projection of $\bar{x}_{s}$ onto $\bar{x}_{0}$. The magnitude of $\bar{x}_{0}$ is the radial distance $r$, so the approximation is

$$
\begin{equation*}
\left|\bar{x}_{s}-\bar{x}_{0}\right| \approx\left|\bar{x}_{0}\right|-\bar{x}_{s} \cdot \frac{\bar{x}_{0}}{\left|\bar{x}_{0}\right|}=r-\bar{x}_{s} \cdot \bar{e}_{r} \tag{8.2.9}
\end{equation*}
$$

The location of the field point is described in Fig. 8.3 in terms of spherical coordinates. The radial unit vector is

$$
\begin{equation*}
\bar{e}_{r}=\sin \psi \cos \theta \bar{e}_{x}+\sin \psi \sin \theta \bar{e}_{y}+\cos \psi \bar{e}_{z} \tag{8.2.10}
\end{equation*}
$$

The location of the surface point is $\bar{x}_{s}=x_{s} \bar{e}_{x}+y_{s} \bar{e}_{y}$, so we have

$$
\begin{equation*}
\left|\bar{x}_{s}-\bar{x}_{0}\right| \approx r-x_{s} \sin \psi \cos \theta-y_{s} \sin \psi \sin \theta \tag{8.2.11}
\end{equation*}
$$

Fig. 8.3 Locations of a surface point $\bar{x}_{s}$ and a field point $\bar{x}_{0}$ for implementation of the farfield approximation of the Rayleigh integral


The condition that $\bar{x}_{0}$ be a farfield point requires that $r \gg x_{s}$ and $y_{s}$ for any point in the surface. In that case $\left|\bar{x}_{s}-\bar{x}_{0}\right|$ in the denominator is well approximated as $1 / r$. The normal velocity is $V_{s}\left(x_{s}, y_{s}\right)$, and an area element is $d x_{s} d y_{s}$. The result is that Eq. (8.1.7) becomes

$$
\begin{equation*}
P_{\mathrm{ff}}\left(\bar{x}_{0}\right)=\frac{i \omega \rho_{0}}{2 \pi r} e^{-i k r} \iint_{\mathcal{S}_{v}} V_{s}\left(x_{s}, y_{s}\right) e^{i k\left(x_{s} \sin \psi \cos \theta+y_{s} \sin \psi \sin \theta\right)} d x_{s} d y_{s} \tag{8.2.12}
\end{equation*}
$$

Suppose that $\mathcal{S}_{v}$ is a rectangle defined by $-a \leq x_{s} \leq a,-b \leq y_{s} \leq b$, and that the velocity distribution has the special form of a product of functions of $x_{s}$ and $y_{s}$, that is, $V_{s}\left(\bar{x}_{s}\right)=X\left(x_{s}\right) Y\left(y_{s}\right)$. Because $d \mathcal{S}=d x_{s} d y_{s}$ and the integration limits are constants, the integrals over $x_{s}$ and $y_{s}$ may be evaluated independently. In this case, the farfield pressure is given by

$$
\begin{equation*}
P_{\mathrm{ff}}\left(\bar{x}_{0}\right)=\frac{i \omega \rho_{0}}{2 \pi r} e^{-i k r} \int_{-a}^{a} X\left(x_{s}\right) e^{i k x_{s} \sin \psi \cos \theta} d x_{s} \int_{-b}^{b} Y\left(y_{s}\right) e^{i k y_{s} \sin \psi \sin \theta} d y_{s} \tag{8.2.13}
\end{equation*}
$$

Each integral fits the definition of a Fourier transform, which is defined in Eq. (B.1.5). The factors appearing in the exponent in each integral are the wavenumbers $\mu_{x}=$ $k \sin \psi \cos \theta$ and $\mu_{y}=k \sin \psi \sin \theta$. The angular spectrum $\tilde{F}\left(\mu_{x}, \mu_{y}\right)$ is defined to be the product of the individual transforms, according to

$$
\begin{equation*}
\tilde{F}\left(\mu_{x}, \mu_{y}\right)=\int_{-a}^{a} X\left(x_{s}\right) e^{i \mu_{x} x_{s}} d x_{s} \int_{-b}^{b} Y\left(y_{s}\right) e^{i \mu_{y} y_{s}} d y_{s} \tag{8.2.14}
\end{equation*}
$$

In terms of this function, the farfield pressure is

$$
\begin{equation*}
P_{\mathrm{ff}}\left(\bar{x}_{0}\right)=i \omega \rho_{0} \frac{e^{-i k r}}{2 \pi r} \tilde{F}(k \sin \psi \cos \theta, k \sin \psi \sin \theta) \tag{8.2.15}
\end{equation*}
$$

If $\mathcal{S}_{v}$ is not rectangular, or if $V_{s}$ is not a product of $x_{s}$ and $y_{s}$ functions, the area integral does not factorize. Nevertheless, integration over the full range of $x_{s}$ and $y_{s}$ leaves only the wavenumbers as variables. The angular spectrum in this case is a surface integral that constitutes a two-dimensional Fourier transform. Thus, Eq. (8.2.15) is descriptive of the farfield pressure in any situation, with the angular spectrum defined to be

$$
\begin{equation*}
\tilde{F}\left(\mu_{x}, \mu_{y}\right)=\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} V_{s}\left(x_{s}, y_{s}\right) e^{i\left(\mu_{x} x_{s}+\mu_{y} y_{s}\right)} d x_{s} d y_{s} \tag{8.2.16}
\end{equation*}
$$

In the case where the angular spectrum factorizes as in Eq. (8.2.14), tabulations of functions and their Fourier transforms are available, see Tables B. 1 and B. 2 in Appendix B. FFT technology may be employed to handle truly complicated velocity functions. However, the normal velocity typically is characterized by an elementary function, which makes it feasible to evaluate the integral analytically. Situations in which the vibrating portion of $\mathcal{S}_{v}$ is circular may be analyzed by using polar coordinates. Such a formulation is the topic of the next section.

Identification that the farfield pressure is directly related to the two-dimensional Fourier transform of the surface velocity leads to an interesting interpretation. Suppose we wish to determine the pressure at a specific field point in the farfield. The radial unit vector to that point is $\bar{e}_{r}$, which is defined in Eq. (8.2.10). The wavenumber of the farfield signal propagating in this direction is $k \bar{e}_{r}$. The trace of this vector onto the $x y$ plane is its projection, $\bar{k}_{s}=k \sin \psi \cos \theta \bar{e}_{x}+k \sin \psi \sin \theta \bar{e}_{y}$ in Fig. 8.4. These components are the wavenumbers $\mu_{x}$ and $\mu_{y}$ for the angular spectrum in Eq. (8.2.15).

This construction makes it evident that $\left|\vec{k}_{s}\right| \leq k$, which leads to an interpretation of the farfield radiation. We construct a surface $\tilde{F}\left(\mu_{x}, \mu_{y}\right)$ in a Cartesian wavenumber


Fig. 8.4 Trace wavenumber vector $\bar{k}_{s}$ for the farfield spherical wave in a designated direction $\bar{e}_{r}$
space such that the height of the surface is the value of $\tilde{F}$, as is done in Fig. 8.4. In the $\mu_{x} \mu_{y}$ plane, we construct a circle of radius $k$. Any point within this circle corresponds to a signal that propagates to the farfield at polar angle $\psi$ and azimuthal angle $\theta$. These angles are related to the wavenumbers $\mu_{x}$ and $\mu_{y}$ of the selected point by $\cos \theta=\mu_{x} /\left(\mu_{x}^{2}+\mu_{y}^{2}\right)^{1 / 2}, \sin \theta=\mu_{y} /\left(\mu_{x}^{2}+\mu_{y}^{2}\right)^{1 / 2}$, and $\cos \psi=\left(\mu_{x}^{2}+\mu_{y}^{2}\right)^{1 / 2} / k$. Selection of a point in the $\mu_{x} \mu_{y}$ plane for which $\left|\bar{k}_{s}\right|>k$ corresponds to subsonic surface waves that do not radiate to the farfield.

This interpretation of the relationship between the surface wave spectrum and the signal that propagates in a specified direction can be used as a design tool. Suppose we wish that the radiated pressure in a certain direction designated by spherical angle $\psi_{0}$ and $\theta_{0}$ be the maximum. For a specified frequency, we can form the wavenumber vector in this direction. The components of this vector in the $x$ and $y$ directions are the corresponding wavenumbers, $\mu_{x}=k \sin \psi_{0} \cos \theta_{0}, \mu_{y}=$ $k \sin \psi_{0} \cos \theta_{0}$. The maximum condition is attained if the two-dimensional Fourier transform of $\bar{n}\left(\bar{x}_{s}\right) \cdot V\left(\bar{x}_{s}\right)$ has a maximum magnitude at this pair of wavenumbers. Finding a surface vibration pattern whose Fourier transform has this attribute would not be a trivial task, but tabulations of Fourier transform pairs like those in Appendix $B$ would be useful.

EXAMPLE 8.1 An electrostatic speaker essentially is a pair of oppositely charged plates, one of which is fixed. The free plate vibrates in response to oscillations of the voltage difference. This vibrating plate may be taken to be rigid, so its normal velocity is $v_{0} \cos (\omega t)$ everywhere on its surface $\mathcal{S}_{v}$. The plate is surrounded on all sides by a large rigid baffle. Two alternative designs are contemplated. Both have equal surface area, but one has a square face whose sides are 200 mm , and the other is 100 mm by 400 mm . The frequency is 2 kHz . Determine the farfield pressure as a function of the polar angles $\psi$ and $\theta$ for each design, and graph $\left(r / \sqrt{S_{v}}\right)\left|P_{\mathrm{ff}}\right| /\left(\rho_{0} c v_{0}\right)$ as a function of the spherical angles.

## Significance

The Fourier transform of the vibration pattern will be found with little effort, which will allow us to focus on the correlation between the properties of the vibration and those of the pressure field.

## Solution

The shape of $\mathcal{S}_{v}$ is rectangular and the velocity function is a constant, so we shall evaluate the angular spectrum directly. Let $2 a$ and $2 b$ be the lengths of the edges parallel to the $x$ and $y$ axes, respectively. Then the complex vibration amplitude is

$$
V_{s}\left(x_{s}, y_{s}\right)=\left\{\begin{array}{l}
v_{0} \text { if }\left|x_{s}\right| \leq a \text { and }\left|y_{s}\right| \leq b \\
0 \text { otherwise }
\end{array}\right.
$$

The two-dimensional Fourier transform of this function is

$$
\begin{aligned}
\tilde{F}\left(\mu_{x}, \mu_{y}\right) & =\int_{-a}^{a} \int_{-b}^{b} v_{0} e^{i\left(\mu_{x} x+\mu_{y} y\right)} d y_{s} d x_{s} \\
& =4 v_{0}\left(\frac{\sin \left(\mu_{x} a\right)}{\mu_{x}}\right)\left(\frac{\sin \left(\mu_{y} b\right)}{\mu_{y}}\right)
\end{aligned}
$$

A compact way to write this expression is to use the sinc function, which is defined as $\operatorname{sinc}(\xi) \equiv \sin (\pi \xi) /(\pi \xi)$. Figure 1 is a plot of the sinc function.


Figure 1.

At $\mu_{x}=0$ or $\mu_{y}=0$, the value of $\tilde{F}\left(\mu_{x}, \mu_{y}\right)$ is the finite value obtained as a limit, which gives $\operatorname{sinc}(0)=1$. The transform of the surface velocity is thereby found to be

$$
\tilde{F}\left(\mu_{x}, \mu_{y}\right)=4 v_{0} a b \operatorname{sinc}\left(\frac{\mu_{x} a}{\pi}\right) \operatorname{sinc}\left(\frac{\mu_{y} b}{\pi}\right)
$$

To evaluate the farfield radiation pattern, we define a hemispherical mesh formed from meridional semicircles through the $z$-axis along which $\theta$ is a constant value between $-\pi$ and $\pi$, and circles centered on the $z$-axis along which $\psi$ is a constant between zero and $\pi / 2$. The radial direction $\bar{e}_{r}$ at each mesh point is described by Eq. (8.2.10). The trace of $k \bar{e}_{r}$ on the $x y$ plane is $k \sin \psi\left(\cos \theta \bar{e}_{x}+\sin \theta \bar{e}_{y}\right)$, so the surface wavenumbers are $\mu_{x}=k \sin \psi \cos \theta$ and $\mu_{y}=k \sin \psi \sin \theta$. With an eye toward computations, we collect variables as nondimensional groups by multiplying and dividing $\tilde{F}\left(\mu_{x}, \mu_{y}\right)$ by $a b$. Correspondingly, Eq. (8.2.15) gives

$$
P_{\mathrm{ff}}\left(\bar{x}_{0}\right)=i \rho_{0} c v_{0} \frac{2 k a b}{\pi} \frac{e^{-i k r}}{r} \operatorname{sinc}\left(\frac{k a \sin \psi \cos \theta}{\pi}\right) \operatorname{sinc}\left(\frac{k b \sin \psi \sin \theta}{\pi}\right)
$$

The value of $a b$ for the alternative designs is the same, and the maximum value of the sinc function is one, so the alternative design gives similar pressure amplitudes. When $b>a$, the oscillation of $\tilde{F}\left(\mu_{x}, \mu_{y}\right)$ in the $\mu_{y}$ direction will be more rapid than that in the $\mu_{x}$ direction. Maxima of $\tilde{F}\left(\mu_{x}, \mu_{y}\right)$ correspond to the directions in which the farfield pressure is a maximum, and zeros of $\tilde{F}\left(\mu_{x}, \mu_{y}\right)$ correspond to
zero farfield pressure in the corresponding $\bar{e}_{r}$ direction. It follows that the slender rectangle configuration will exhibit more side lobes along $\theta=90^{\circ}$ and $270^{\circ}$ than it does along $\theta=0$ and $180^{\circ}$. To see the details, we need to evaluate the pattern. The area on which the vibration occurs is $4 a b$, so the specified quantity to plot is given by

$$
\left(\frac{r}{2(a b)^{1 / 2}}\right) \frac{\left|P_{\mathrm{ff}}\left(\bar{x}_{0}\right)\right|}{\rho_{0} c v_{0}}=\frac{k(a b)^{1 / 2}}{\pi}\left|\operatorname{sinc}\left(\frac{k a \sin \psi \cos \theta}{\pi}\right) \operatorname{sinc}\left(\frac{k b \sin \psi \sin \theta}{\pi}\right)\right|
$$

Figure 2 describes the square configuration at 2 kHz . The data is displayed as a spherical plot in which the radial distance is the quantity defined above. This is accompanied by polar plots at $\theta=0$, which is the $x z$ plane, and $\theta=90^{\circ}$, which is the $y z$ plane. (The range of $\psi$ was set as $-\pi / 2<\psi<\pi / 2$ to depict the data on both sides of the $z$-axis.) Because $a=b$, the polar plots are identical. The spherical plot is close to being axisymmetric relative to the $z$-axis, and the pressure falls off rapidly as $\psi$ increases from zero. This type of pattern, in which the signal is essentially aimed in a certain direction, is said to be a sound beam. There are several side lobes, but the pressure amplitude in them is quite low relative to the main lobe.


Figure 2. $\left(r /\left(a b^{1 / 2}\right)\right)|P| /\left(\rho_{0} c v_{0}\right)$ at 2 kHz , $a=b=200 \mathrm{~mm}$.

Figure 3 describes the case where the plate is a narrow rectangle at 2 kHz , as in the previous case. This field also is a sound beam. The polar plots indicate that the signal in the $y z$ plane is much more tightly confined to the $z$-axis than it is in the $x z$ plane. This might seem to be counterintuitive because the width of the plate in the $x$ direction is large. To explain this feature, consider the limit in which the plate's dimensions $a$ and $b$ become infinite. In that case, the radiated field would consist of a simple plane wave that radiates in the $z$ direction. Such a wave would be zero in all directions other than $\psi=0$.


Figure 3. $\left(r /\left(a b^{1 / 2}\right)\right)|P| /\left(\rho_{0} c v_{0}\right)$ at 2 kHz , $a=100 \mathrm{~mm}, b=400 \mathrm{~mm}$.

We have encountered a number of situations in which decreasing the size of the vibrating object (or the frequency) produces a field that varies more gradually in space. Figure 4 exhibits this feature. The values of $a$ and $b$ are like those of Fig. 2, but the frequency now is 200 Hz . Even though the rectangular plate is long and narrow (we say that its aspect ratio $b / a$ is high), the field is almost axisymmetric, and there is little variation from $\psi=0$ to $\psi=90^{\circ}$. The nondimensional parameters setting the behavior are $k a$ and $k b$, which are 0.37 and 1.48 , respectively, in this last case. Thus, both sides are much smaller than the acoustic wavelength, $2 \pi / k$. As a general rule, a fairly uniform farfield will be obtained if the size of the radiating body is substantially less than the acoustic wavelength.


Figure 4. $\left(r /\left(a b^{1 / 2}\right)\right)|P| /\left(\rho_{0} c v_{0}\right)$ at 200 Hz ,

$$
a=100 \mathrm{~mm}, b=400 \mathrm{~mm}
$$

### 8.2.2 Farfield of a Piston Transducer

In the previous section, the surface vibration was considered to be best described in terms of a Cartesian coordinate system. Here, we consider the situation in which the outline of the vibrating region is circular. A set of polar coordinates $\left(R_{s}, \theta_{s}\right)$ centered on this region are most suitable to formulate the Rayleigh integral. The development will begin by allowing the velocity distribution to depend arbitrarily on $R_{s}$ and $\theta_{s}$. However, in most cases of interest the vibration pattern is axisymmetric, which means that the vibration amplitude depends solely on $R_{s}$. A special case is a piston transducer, for which the vibration amplitude is uniform across the circular area. This model often is used to represent a common loudspeaker. Some very interesting phenomena are encountered in this case, so it will receive the most attention.

Figure 8.5 depicts a vibrating circular region $S_{v}$ surrounded by a rigid baffle. The $z$-axis for a cylindrical coordinate system is normal to the surface. The field point $\bar{x}_{0}$ is located by its spherical coordinates $(r, \psi, \theta)$, and a generic point on $\mathcal{S}_{v}$ has polar coordinates $\left(R_{s}, \theta_{s}\right)$, so its position is $\bar{x}_{s}=R_{s} \cos \theta_{s} \bar{e}_{x}+R_{s} \sin \theta_{s} \bar{e}_{y}$.

The farfield approximation of the distance between a surface point and a field point is $\left|\bar{x}-\bar{x}_{s}\right|=r-\bar{x}_{s} \cdot \bar{e}_{r}$. In view of Eqs. (8.2.9) and (8.2.10), this reduces to

$$
\begin{equation*}
\left|\bar{x}-\bar{x}_{s}\right|=r-R_{s} \sin \psi \cos \left(\theta_{s}-\theta\right) \tag{8.2.17}
\end{equation*}
$$

In combination with a polar coordinate description of $d \mathcal{S}$, this expression converts the farfield approximation of the Rayleigh integral, Eq. (8.2.12), to

$$
\begin{equation*}
P_{\mathrm{ff}}\left(\bar{x}_{0}\right)=\frac{i \omega \rho_{0}}{2 \pi r} e^{-i k r} \int_{0}^{a} \int_{-\pi}^{\pi} V_{s}\left(R_{s}, \theta_{s}\right) e^{i k R_{s} \sin \psi \cos \left(\theta_{s}-\theta\right)} R_{s} d \theta_{s} d R_{S} \tag{8.2.18}
\end{equation*}
$$

Fig. 8.5 Coordinate descriptions of a surface point and a field point suitable for the analysis of a circular transducer in an infinite baffle

where $a$ is the radius of the area over which the vibration occurs.
Usually the velocity distribution is axisymmetric, so $V_{z}$ is independent of $\theta_{s}$. This greatly simplifies the analysis because it allows the integration over $\theta_{s}$ to be factored out. An integration over $\theta_{s}$ with $R_{s}$ held fixed represents the contribution to the farfield pressure of a circular ring whose radius is $R_{s}$ and whose radial width is $d R_{s}$. The normal velocity along this ring is $V_{s}\left(R_{s}\right)$, so that

$$
\begin{equation*}
d P_{\text {ring }}=\frac{i \omega \rho_{0}}{2 \pi r} e^{-i k r} V_{s}\left(R_{s}\right) R_{s} d R_{s} \int_{-\pi}^{\pi} e^{i k R_{s} \sin \psi \cos \left(\theta_{s}-\theta\right)} d \theta_{s} \tag{8.2.19}
\end{equation*}
$$

The value of $\theta$ is fixed by selection of the field point. This allows us to change the integration variable to $\theta^{\prime}=\theta_{s}-\theta$. Furthermore, the integrand is periodic in any $2 \pi$ interval, and it is an even function of $\theta^{\prime}$. This allows us to integrate over $0 \leq \theta^{\prime} \leq \pi$, then double the result. Hence, the pressure radiated by a ring whose radius is $R_{s}$ and width is $d R_{s}$ is given

$$
\begin{equation*}
d P_{\text {ring }}=\frac{i \omega \rho_{0}}{\pi r} e^{-i k r} V_{s}\left(R_{s}\right) R_{s} d R_{s} \int_{0}^{\pi} e^{i k R_{s} \sin \psi \cos \theta_{s}} d \theta_{s} \tag{8.2.20}
\end{equation*}
$$

The integral may be found in most references on Bessel functions, ${ }^{2}$ which leads to

$$
\begin{equation*}
d P_{\text {ring }}=i \rho_{0} c V_{s}\left(R_{s}\right) k R_{s} d R_{s} \frac{e^{-i k r}}{r} J_{0}\left(k R_{s} \sin \psi\right) \tag{8.2.21}
\end{equation*}
$$

There is no dependence on the azimuthal angle $\theta$ for the field point because any point-symmetric vibration pattern will generate an axisymmetric pressure field.

Equation (8.2.21) may be used for an actual ring by replacing the differential width $d R_{s}$ with the actual width $\Delta R_{s}$ and taking $R_{s}$ to be the mean radius. However, doing so requires that $\Delta R_{s} / R_{s}$ be sufficiently small to that any fluctuation of the $V_{z}$ value from the inner to outer radius is negligible. The more important use is to

[^22]construct the signal for a full circular region. The Rayleigh integral in that case is essentially a sum of contributions of all differential elements, so that
\[

$$
\begin{equation*}
P_{\mathrm{ff}}\left(\bar{x}_{0}\right)=i \rho_{0} c k \frac{e^{-i k r}}{r} \tilde{f}(k \sin \psi), \quad \tilde{f}(\mu)=\int_{0}^{a} V_{s}\left(R_{s}\right) J_{0}\left(\mu R_{s}\right) R_{s} d R_{s} \tag{8.2.22}
\end{equation*}
$$

\]

This expression has the appearance of Eq. (8.2.15). The function $\tilde{f}(\mu)$ is the angular spectrum for a decomposition of the surface velocity into waves that propagate along the surface radially outward from the center. From a mathematical perspective, it is the Hankel transform of the axisymmetric $V_{s}\left(R_{s}\right)$. The Hankel transform may be derived by transforming the two-dimensional Fourier transform to polar coordinates. ${ }^{3}$ Depending on the nature of $V_{s}\left(R_{s}\right)$, it might be possible to evaluate $\tilde{f}(\mu)$ analytically. If not, it can be evaluated numerically.

A particularly important case is that of a piston or translating rigid disk, for which $V_{s}=V_{0}$ for $R_{s} \leq a$. Introduction of $u=k R_{s} \sin \psi$ into the integrand of the angular spectrum leads to a standard integral, ${ }^{4}$ with the eventual result that

$$
\begin{equation*}
P_{\mathrm{ff}}\left(\bar{x}_{0}\right)=i \rho_{0} c V_{0} \frac{J_{1}(k a \sin \psi)}{\sin \psi} \frac{a}{r} e^{-i k r} \tag{8.2.23}
\end{equation*}
$$

The Rayleigh distance $R_{0}$ is a reference length often used to describe acoustic radiation. We will see in the next section that it indicates the radial distance at which the signal is well represented by its farfield approximation. The definition of $R_{0}$ is that it is the ratio of the area over which the vibration occurs to the wavelength. For a circular piston, it is

$$
\begin{equation*}
R_{0}=\frac{\mathcal{S}_{v}}{\lambda}=\frac{\pi a^{2}}{2 \pi / k}=\frac{1}{2} k a^{2} \tag{8.2.24}
\end{equation*}
$$

A simple expression for the maximum pressure at a specified radial distance, which occurs at $\psi=0$, results from using the Rayleigh distance in Eq. (8.2.23). As $\psi \rightarrow 0$, the first order Bessel function approaches $(1 / 2) k a \sin \psi$, so that

$$
\begin{equation*}
\max \left(\left|P_{\mathrm{ff}}\left(\bar{x}_{0}\right)\right|\right)=\rho_{0} c V_{0} \frac{R_{0}}{r} @ \psi=0 \tag{8.2.25}
\end{equation*}
$$

The corresponding farfield pressure is

$$
\begin{equation*}
P_{\mathrm{ff}}\left(\bar{x}_{0}\right)=i \rho_{0} c V_{0}\left(\frac{2 J_{1}(k a \sin \psi)}{k a \sin \psi}\right)\left(\frac{R_{0}}{r}\right) e^{-i k r} \tag{8.2.26}
\end{equation*}
$$

[^23]The directivity is the ratio of the pressure magnitude as a function of $\psi$ to the value on-axis, with $r$ fixed in the farfield. Thus,

$$
\begin{equation*}
\mathcal{D}(\psi)=\left.\frac{\left|P_{\mathrm{ff}}\right|}{\max \left(\left|P_{\mathrm{ff}}\right|\right)}\right|_{r>R_{0}}=\frac{2\left|J_{1}(k a \sin \psi)\right|}{k a \sin \psi} \tag{8.2.27}
\end{equation*}
$$

An interesting comparison is with the pressure radiated by a ring having the same radius $a$ and vibrational velocity $V_{0}$. We find from Eq. (8.2.21) that the ratio $\left(P_{\mathrm{ff}}\right)_{\text {ring }} /\left(P_{\mathrm{ff}}\right)_{\text {piston }}=2 \Delta R_{S} / a$ on the $z$-axis, where $\psi=0$. The expression for a ring is based on the width $\Delta R_{s}$ being much less than the radius, from which it follows that the ring radiates much more weakly as a consequence of the smaller face area. The directivity factor for a ring is $J_{0}(k a \sin \psi)$. The nulls of the directivity in each case correspond to parameters for which $k a \sin \psi$ is a zero of the respective Bessel functions.

The number of nulls that occurs is limited by the requirement that $\sin \psi<1$, from which it follows that this number increases in a stepwise manner approximately in proportion to the value of $k a$. Thus, below $k a=3.832$ for a piston and $k a=2.405$ for a ring, the directivity exhibits no nulls. Above these frequencies, the interval from the axis to the first null is the main lobe, and the other intervals are the side lobes. The maxima in the side lobes decrease rapidly with increasing angle from the $z$-axis. Figure 8.6 displays some typical patterns. The field in each case is axisymmetric, so a polar plot over $-\pi / 2 \leq \psi \leq \pi / 2$ is sufficient.

The rapid decrease in the magnitude of the side lobes obscures them in a polar plot. A rectangular plot, such as that in Fig. 8.7, conveys a clearer quantitative picture.


Fig. 8.6 Directivity of a piston and a ring having radius $a$ embedded in an infinite baffle


Fig. 8.7 Cartesian plot of the directivity of a piston and a ring embedded in an infinite baffle, $k a=18$

We can discern some general trends from Fig. 8.7 and Table 8.1. For a given value of $k a$, the main lobe is narrower with a ring than it is with a piston. The ring has more side lobes, and the side lobes are stronger. These trends, in combination with the stronger signal radiated by the piston, indicate that if one's objective is to aim a signal in a certain direction at a specified frequency, they should use the piston configuration. The piston's radius should be as large as feasible, because a large $k a$ will yield a small value of $\psi$ for the first null.

Table 8.1 Polar angles at which the directivity of a piston and a ring have a null

|  | $k a \sin \psi$ |  |  |  |
| :--- | :--- | :--- | :--- | :--- |
| Null \# | 1 | 2 | 3 | 4 |
| Piston | 3.832 | 7.016 | 10.173 | 13.324 |
| Ring | 2.405 | 5.520 | 8.654 | 11.792 |

A quantity of interest is the beamwidth $\Delta_{\mathrm{bw}}$, which is the cone angle at which the intensity is half the maximum on-axis value. In the farfield, the radial particle velocity is $p /\left(\rho_{0} c\right)$ and the other velocity components are negligible, so the time-averaged intensity is

$$
\begin{equation*}
\left(I_{r}\right)_{\mathrm{av}}=\frac{\left|P^{2}\right|}{2 \rho_{0} c}=\frac{\rho_{0} c\left|V_{0}\right|^{2}}{2}\left(\frac{R_{0}}{r}\right) \mathcal{D}(\psi)^{2} \tag{8.2.28}
\end{equation*}
$$

The half-power point corresponds to $\mathcal{D}(\psi)=1 / \sqrt{2}$. For the piston directivity in Eq. (8.2.27), this condition occurs at $k a \sin \psi=1.62$, whereas Table 8.1 indicates that the first null occurs at $k a \sin \psi=3.82$. Hence, if large $k a$ is large, the beamwidth approximately subtends the center half of the main lobe.

A question that sometimes is posed is why is the piston model used to describe the radiation properties of transducers that do not look like a piston? One such device is a speaker cone embedded in a planar baffle, which is a common driver for sound reproduction systems. This arrangement is shown in Fig. 8.8.

The field is symmetric about the $z$-axis. The depth of the transducer below the $x y$ plane of the baffle is some function $Z(R)$. A well-designed speaker cone translates in the $z$ direction, so that all points on the cone have the same velocity

Fig. 8.8 Cross-sectional view of a conical speaker embedded in a rigid baffle

$\operatorname{Re}\left(V_{z} e^{i \omega t}\right)$. Because the angle of the cone is close to $90^{\circ}$, the normal to the cone's surface is nearly parallel to the $z$-axis, so we may say that the (time domain) surface velocity is

$$
n\left(\bar{x}_{s}\right) \cdot \bar{v}\left(\bar{x}_{s}\right)=\left\{\begin{array}{l}
\operatorname{Re}\left(V_{z} e^{i \omega t}\right) @ z=-Z(R) \text { if }|R|<a  \tag{8.2.29}\\
0 \text { if }|R|>a
\end{array}\right.
$$

An approximation of the field close to the surface is that each point emits a plane wave that propagates in the axial direction. (This is a crude representation of a complicated process, but it is adequate for the present purpose.) In this approximation, a signal that is emitted from a point on the transducer's face at radial distance $R$ requires an interval of $Z(R) / c$ to arrive at the $x y$ plane. This is a lag, so the phase delay for a harmonic signal at frequency $\omega$ is $k Z(R)$. The complex surface velocity on the $x y$ plane in this approximation is

$$
V_{s}=\left\{\begin{array}{l}
V_{z} e^{-i k Z(R)} \text { if }|R|<a  \tag{8.2.30}\\
0 \text { if }|R|>a
\end{array}\right.
$$

In most cases, the intent is to radiate sound over a broad range of polar angles, in which case $k a$ will be small. The preceding description of $V_{s}$ states that there is a position-dependent phase lag, which may be written as $k a(Z(R) / a)$. If $k a$ is not much larger than one and the cone angle is close to $90^{\circ}$, so that $Z(R) / a$ is small everywhere, then it is reasonable to ignore this phase lag. A different way of viewing this is to write the phase lag as $2 \pi Z(R) / \lambda$, where $\lambda$ is the wavelength of a plane wave at frequency $\omega$. In this view, a nonplanar piston can be considered to behave like a piston if its maximum depth is much less than the plane wavelength. Conversely, in many cases, especially ultrasonic applications like acoustical microscopy, a transducer is driven at a very large $k a$ in order to obtain a narrow main lobe. In such cases, it would be wrong to use a piston model to analyze the radiated field. Equation (8.2.30) will be used in Example 8.4 to develop an approximate nonpiston model for the pressure radiated by a nonplanar transducer at high $k a$.

EXAMPLE 8.2 Two designs for a circular transducer embedded in a rigid baffle are under consideration. One is a conventional piston, whose surface velocity is $v_{0} \cos (\omega t)$ everywhere on its face. The face of the other is a mem-
brane that is stretched across the circumference. Its surface velocity is well approximated be a parabolic distribution, that is, $V_{s}=v_{1}\left(1-R^{2} / a^{2}\right)$. The medium in which these devices will operate is water. It is desired that the pressure amplitude at a distance of 100 m for either transducer be 0.8 atm , and that the first null of the farfield directivity be at $\psi=2^{\circ}$ in either case. The frequency at which the transducers will operate is 30 kHz in order to test the hearing of sea mammals. For each design, determine the required radius and the maximum velocity on its face. Then, determine the polar angle of its second null and the maximum pressure in the side lobe between the first and second nulls.

## Significance

Design problems like this improve our understanding of the relationship between system parameters and performance.

## Solution

The first task is to determine the directivity associated with the parabolic velocity profile. The farfield description in Eq. (8.2.22) for the given velocity distribution is simplified slightly by changing the integration variable to $\xi=R_{s} / a$, which gives

$$
\begin{equation*}
P_{\mathrm{ff}}\left(\bar{x}_{0}\right)=i \rho_{0} c v_{1} k a^{2} \frac{e^{-i k r}}{r} \int_{0}^{1}\left(1-\xi^{2}\right) J_{0}(k a \xi \sin \psi) \xi d \xi \tag{1}
\end{equation*}
$$

We could evaluate the integral numerically for each value of $k a \xi \sin \psi$, but an analytical evaluation is possible. A tabulated integral ${ }^{5}$ states that

$$
\begin{align*}
& \int_{0}^{\pi / 2} J_{\mu}(z \sin t)(\sin t)^{\mu+1}(\cos t)^{2 \nu+1} d t  \tag{2}\\
& =\frac{2^{\nu} \Gamma(\nu+1)}{z^{\nu+1}} J_{\mu+\nu+1}(z), \quad \operatorname{Re}(\mu)>-1 \text { and } \operatorname{Re}(\nu)>-1
\end{align*}
$$

where $\Gamma$ denotes the gamma function. Our interest is with integer values, in which case $\Gamma(\nu+1) \equiv \nu$ ! If this formula is to apply to Eq. (1), then the order of the Bessel functions and the argument should match. This suggests that we should set $\mu=0$, $\sin t=\xi$, and $z=k a \sin \psi$. Furthermore, setting $\sin t=\xi$ leads to matching integration limits. This change of variables gives $(\cos t) d t=d \xi$ and $(\cos t)^{2 \nu}=\left(1-\xi^{2}\right)^{\nu}$. Thus, setting $\xi=\sin t, k a \sin \psi=z, \mu=0$, and $\nu=1$ transforms Eq. (1) to the integral in Eq. (2). The result is that the farfield for the membrane transducer is

$$
\begin{equation*}
P_{\mathrm{ff}}\left(\bar{x}_{0}\right)=i \rho_{0} c v_{1} k a^{2}\left[\frac{2 J_{2}(k a \sin \psi)}{(k a \sin \psi)^{2}}\right] \frac{e^{-i k r}}{r} \tag{3}
\end{equation*}
$$

[^24]As we did for the piston in Eq. (8.2.27), we define the directivity relative to the maximum farfield pressure. At $\psi=0$, the denominator is singular, but the asymptotic expansion of the Bessel function gives $J_{2}(k a \sin \psi) \approx(k a \sin \psi)^{2} / 8$, so we find that

$$
\begin{gather*}
\max \left(\left|P_{\mathrm{ff}}\left(\bar{x}_{0}\right)\right|\right)=\rho_{0} c v_{1} \frac{a}{r}\left(\frac{k a}{4}\right) \equiv \frac{1}{2} \rho_{0} c v_{1} \frac{R_{0}}{r} @ \psi=0 \\
\mathcal{D}(\psi)=\frac{\left|P_{\mathrm{ff}}\left(\bar{x}_{0}\right)\right|}{\max \left(\left|P_{\mathrm{ff}}\left(\bar{x}_{0}\right)\right|\right)}=\frac{8\left|J_{2}(k a \sin \psi)\right|}{(k a \sin \psi)^{2}} \tag{4}
\end{gather*}
$$

Before we address the design criteria, it is worth noting that if $v_{1}$ and $k a$ for the membrane equal $v_{0}$ and $k a$ for the piston, then the maximum farfield pressure in Eq. (4) is half the value for a piston in Eq. (8.2.26). Separately, we know that at very low frequencies the radiated pressure is omnidirectional, with an amplitude that is proportional to the complex volume velocity amplitude. This value for each design is

$$
\begin{gathered}
\left(\hat{Q}_{s}\right)_{\text {piston }}=\iint_{\mathcal{S}} V_{s} d S=v_{0}\left(\pi a^{2}\right) \\
\left(\hat{Q}_{s}\right)_{\text {parabolic }}=\iint_{\mathcal{S}} v_{1}\left(1-\frac{R_{s}^{2}}{a^{2}}\right) 2 \pi R_{s} d R_{s}=v_{1}\left(\frac{1}{2} \pi a^{2}\right)
\end{gathered}
$$

Thus, the ratio of maximum farfield pressures for the two designs is the same as the ratio of volume velocities.

One design criterion is the value of $\psi$ for the first null, which corresponds to the first positive zero of the respective Bessel function. For $J_{1}(k a \sin \psi)$, we have Table 8.1. We may find the zeros of $J_{2}(k a \sin \psi)$ by inspection of a graph of $J_{2}(x)$, and a high precision value may be obtained by using numerical methods to find the roots of $J_{2}(x)=0$. The first few values are

$$
J_{2}(x)=0 @ x=5.136,8.417,11.620,14.796
$$

The frequency is $\omega=60,000 \pi \mathrm{rad} / \mathrm{s}$, and we take $c=1480 \mathrm{~m} / \mathrm{s}$. For a first null at $\psi=2^{\circ}$, matching $k a \sin \psi$ to the first zero of the respective functions gives

$$
\begin{gathered}
\text { Piston: } k a=\frac{3.832}{\sin \left(2^{\circ}\right)}=109.80 \Longrightarrow a=0.862 \mathrm{~m} \\
\text { Membrane: } k a=\frac{5.136}{\sin \left(2^{\circ}\right)}=147.15 \Longrightarrow a=1.155 \mathrm{~m}
\end{gathered}
$$

Both are enormous designs, but such transducers can be found on some ocean-going research vessels.

The other design criterion sets the maximum pressure at 100 m to $0.8 \mathrm{~atm}=$ 81.06 kPa . The Rayleigh distance for the membrane transducer, which is the larger value, is $R_{0}=85 \mathrm{~m}$, so $r=100$ for the specified pressure is in the farfield. Matching the maximum farfield pressure to the specified value in each case gives

Piston: $\rho_{0} c v_{1} \frac{a}{r}\left(\frac{k a}{4}\right)=1000(1480) v_{0}\left(\frac{0.862}{100}\right)\left(\frac{24.70}{2}\right)=81\left(10^{3}\right) \Longrightarrow$

$$
v_{0}=0.1156 \mathrm{~m} / \mathrm{s}
$$

Membrane: $\rho_{0} c v_{1} \frac{a}{r}\left(\frac{k a}{4}\right)=1000(1480) v_{1}\left(\frac{1.155}{100}\right)\left(\frac{33.11}{4}\right)=81\left(10^{3}\right) \Longrightarrow$

$$
\begin{equation*}
v_{1}=0.1287 \mathrm{~m} / \mathrm{s} \tag{8.2.31}
\end{equation*}
$$

The required maximum surface velocities for the transducers are close, because the membrane transducer has a larger radius, which compensates for its smaller volume velocity per unit surface area.

The last factor to consider pertains to the first side lobe, which is bounded by the first and second nulls. The second null is located at $k a \sin \psi=7.016$ for the piston, and $k a \sin \psi=8.417$ for the membrane, which corresponds to $\psi=3.66^{\circ}$ and $\psi=3.28^{\circ}$, respectively. We could perform a mathematical search for the maximum in each side lobe, but it is easier to scan the data for a plot of each directivity. The angular dependence for both configurations are depicted in Fig. 1. The first side lobe maxima are $\mathcal{D}=0.1323$ at $\psi=2.69^{\circ}$ for the piston and $\mathcal{D}=0.0586$ at $\psi=2.49^{\circ}$ for the membrane transducer. The graphs show that although the main lobes are closely matched, the side lobes of the membrane transducer are narrower and much lower in amplitude. These differences are consequences of the fact that $v_{s}$ is a continuous function of $R_{s}$ for the membrane transducer, whereas it is discontinuous at the baffle for the piston. Discontinuities generally lead to the occurrence of diffraction effects.


Figure 1.

### 8.3 Axial Dependence for a Circular Transducer

The farfield approximation yields valuable information, but it does not provide any insight regarding the range at which it is applicable. The analysis in this section evaluates the Rayleigh integral for any distance from the surface. However, the analysis that follows is not general. It requires that the vibrating region of the surface be circular, that the pattern of the surface vibration be axisymmetric, and that the field point be situated on the axis of symmetry. One of its benefits is that the result will shed light on the nearfield properties, and how those properties transition to the farfield.

The description of the location of a surface point and a field point in Fig. 8.5 are suitable here. The field point is on the $z$-axis, so $\psi=0$. Then, the distance from a surface point to a field point is $\left|\bar{x}_{0}-\bar{x}_{s}\right|=\left(z^{2}+R_{s}^{2}\right)^{1 / 2}$. As a consequence of taking the surface vibration to be axisymmetric, the surface velocity is $V_{s}\left(R_{s}\right)$. Furthermore, because $\left|\bar{x}-\bar{x}_{s}\right|$ does not depend on the polar angle, we may use a ring of radius $R_{s}$ and width $d R_{s}$ as the differential element of area, so $d \mathcal{S}=2 \pi R_{s} s R_{s}$. According to Eq. (8.1.7), the contribution of this ring to the pressure at the field point is

$$
\begin{equation*}
d P_{\mathrm{axial}}(z)=i \omega \rho_{0} V_{s}\left(R_{s}\right) \frac{e^{-i k\left(z^{2}+R_{s}^{2}\right)^{1 / 2}}}{\left(z^{2}+R_{s}^{2}\right)^{1 / 2}} R_{s} d R_{s} \tag{8.3.1}
\end{equation*}
$$

This differential contribution is integrated over the surface area, where the vibration occurs. Thus, we find that

$$
\begin{equation*}
P_{\text {axial }}(z)=i \omega \rho_{0} \int_{0}^{a} V_{s}\left(R_{s}\right) \frac{e^{-i k\left(z^{2}+R_{s}^{2}\right)^{1 / 2}}}{\left(z^{2}+R_{s}^{2}\right)^{1 / 2}} R_{s} d R_{s} \tag{8.3.2}
\end{equation*}
$$

In most situations, evaluation of this integral will require numerical methods, but an analytical integration is possible in the special case of a translating piston, for which $V_{s}\left(R_{s}\right)=V_{0}$. Toward that end, we change the integration variable from $R_{s}$ to the radial distance $r$ from a point on a ring to the field point. Because $r^{2}=R_{s}^{2}+z^{2}$ and $z$ is constant in the integral, differentiation of this relation gives $r d r=R_{s} d R_{s}$. Thus, the change of variables leads to

$$
\begin{equation*}
P_{\text {axial }}(z)=i \omega \rho_{0} V_{0} \int_{z}^{\left(z^{2}+a^{2}\right)^{1 / 2}} e^{-i k r} d r \tag{8.3.3}
\end{equation*}
$$

This change of variables gives the integrand the appearance of a plane wave, in that there is no $1 / r$ term to account for spreading loss. This perception seems to be borne out by the integrated form, which is

$$
\begin{equation*}
P_{\text {axial }}(z)=\rho_{0} c V_{0}\left(e^{-i k z}-e^{-i k\left(z^{2}+a^{2}\right)^{1 / 2}}\right) \tag{8.3.4}
\end{equation*}
$$

The first term is the center wave because its phase is that of a wave that propagates distance $z$ from the center of the piston to the field point. The distance from the edge
of the piston to the field point is $\left(z^{2}+a^{2}\right)^{1 / 2}$, so the second term is the edge wave. In this view, Eq. (8.3.3) represents a superposition of waves from rings that begin at the center and end at the edge. The integrated form says that the contributions of all the ring waves between the center and edge annihilate each other, but there is nothing coming from beyond $R_{s}=a$, which leaves the edge wave as the cancelling effect.

A clearer picture is obtained by using the identities that

$$
\begin{gather*}
z=\frac{1}{2}\left[z+\left(z^{2}+a^{2}\right)^{1 / 2}\right]+\frac{1}{2}\left[z-\left(z^{2}+a^{2}\right)^{1 / 2}\right] \\
\left(z^{2}+a^{2}\right)^{1 / 2}=\frac{1}{2}\left[z+\left(z^{2}+a^{2}\right)^{1 / 2}\right]-\frac{1}{2}\left[z-\left(z^{2}+a^{2}\right)^{1 / 2}\right] \tag{8.3.5}
\end{gather*}
$$

Substitution of these expressions into Eq. (8.3.4) leads to

$$
\begin{equation*}
P_{\text {axial }}(z)=2 i \rho_{0} c V_{0} e^{-i(k / 2)\left[z+\left(z^{2}+a^{2}\right)^{1 / 2}\right]} \sin \left[\frac{k}{2}\left(\left(z^{2}+a^{2}\right)^{1 / 2}-z\right)\right] \tag{8.3.6}
\end{equation*}
$$

The complex exponential factor gives the phase lag relative to the surface velocity. This lag does not increase linearly with increasing $z$. The sine term is more important because it tells us that the magnitude of the pressure on the $z$-axis is

$$
\begin{equation*}
\left|P_{\text {axial }}(z)\right|=2 \rho_{0} c\left|V_{0}\right|\left|\sin \left(\frac{k\left(a^{2}+z^{2}\right)^{1 / 2}-k z}{2}\right)\right| \tag{8.3.7}
\end{equation*}
$$

The maximum pressure that is observed on the axis of symmetry is $2 \rho_{0} c V_{0}$, which is twice the plane wave value. Maxima occur at discrete locations called antinodes that are separated by nodes at which the pressure is zero. Although the function describing this oscillation is sinusoidal, the argument of the function is not proportional to $z$, so the nodes and antinodes are not spaced evenly along the $z$-axis. Nodes occur at locations where the argument is $n \pi$, with $n>0$ because $z \geq 0$. This condition corresponds to

$$
\begin{equation*}
\text { Nodes: } k z=\frac{(k a)^{2}-(2 n \pi)^{2}}{4 n \pi}, n=1,2, \ldots, \text { floor }\left(\frac{k a}{2 \pi}\right) \tag{8.3.8}
\end{equation*}
$$

For antinodes, the argument of the sine function must be $(2 n-1) \pi / 2$, also with $n>0$. This leads to

$$
\begin{equation*}
\text { Antinodes: } k z=\frac{(k a)^{2}-(2 n-1)^{2} \pi^{2}}{2(2 n-1) \pi}, \quad n=1,2, \ldots, \text { floor }\left(\frac{k a+\pi}{2 \pi}\right) \tag{8.3.9}
\end{equation*}
$$

The corollary of these relations is that nodes do not exist if $k a<2 \pi$, and antinodes do not exist if $k a<\pi$. Below $k a=\pi$, the pressure decreases monotonically with increasing $z$. The pressure amplitude at $z=0$, which marks the center of the piston, is found from Eq. (8.3.7) to be $2 \rho_{0} c V_{0} \sin (k a / 2)$ ). These observations lead
us to recognize that decreasing the frequency below $k a=\pi$ will have the effect of decreasing the pressure everywhere along the axis of symmetry. Above $k a=2 \pi$, the farthest antinode ( $n=1$ ) is always more distant than the farthest node.

An important aspect is the behavior at large $z$. A binomial expansion of Eq. (8.3.6) based on $a / z$ being small gives

$$
\begin{equation*}
P_{\text {axial }}(z)=2 i \rho_{0} c V_{0} e^{-i k z} \sin \left[\frac{k z}{2}\left(1+\frac{a^{2}}{z^{2}}\right)^{1 / 2}-\frac{k z}{2}\right] \approx 2 \rho_{0} c V_{0} e^{-i k z} \sin \left(\frac{R_{0}}{2 z}\right) \tag{8.3.10}
\end{equation*}
$$

where $R_{0}$ is the Rayleigh distance defined in Eq. (8.2.24). This expression is valid whenever $z \gg a$. If $z$ is sufficiently greater than $R_{0}$, then the sine function may be approximated by its argument. As it must to be consistent with the general property of spherical spreading in the farfield, the pressure amplitude tends to a $1 / z$ dependence,

$$
\begin{equation*}
P_{\text {axial }}(z)_{\mathrm{ff}}=\rho_{0} c V_{0} \frac{R_{0}}{z} e^{-i k z} \tag{8.3.11}
\end{equation*}
$$

When we set $\psi=0$ in Eq. (8.2.27), so that $r=z$, the result is identical to this expression. This equivalence suggests that $r>R_{0}$ marks the beginning of the farfield. Although this suggestion is based on the axial behavior, it is confirmed by mappings of the field obtained from numerical evaluations of the Rayleigh integral, as we will see in the next section.

Figure 8.9, for $k a=30$, is a typical axial pattern. There are four nodes and five antinodes. The spacing between these locations increases with increasing axial distance, with the increment to the last antinode being much greater than the others. The antinodes occur sharply, but peaks also occur over distances that are a small fraction of the piston radius. This is an important feature if one wishes to measure the nearfield at high frequencies. It means that the microphone/hydrophone must have a very small diameter in order to probe these small regions, and the position must be precisely controlled. Figure 8.9 also shows that the farfield approximation is quite good from the Rayleigh distance outward.


Fig. 8.9 Dependence on distance along the axis of symmetry of the pressure radiated by a piston in an infinite baffle, $k a=30$

EXAMPLE 8.3 The chief engineer of Weird Acoustical Devices, Inc. has the idea to make an "anti-piston" transducer to be embedded in a baffle. The idea is to use piezoceramic elements that form an inner circle that covers $0 \leq R_{s}<a / 2$, and a ring covering $a / 2 \leq R_{s}<a$. Each set of elements will be driven by a separate circuit in order that the surface velocity on the outer ring will have the same magnitude but $180^{\circ}$ out-of-phase relative to the inner ring's surface velocity. Derive an expression for the axial pressure radiated by this device. Compare it to the axial field of a conventional piston transducer at $k a=10$ and 40. In particular, compare the pattern of nodes and antinodes and the distance at which the farfield approximation becomes accurate.

## Significance

This example will greatly enhance our understanding of the use of the Rayleigh integral to evaluate an axial field, and the results will highlight the significance of center and edge waves.

## Solution

The given surface velocity is

$$
V_{s}=\left\{\begin{array}{l}
V_{0} \text { if } 0 \leq R_{s}<a / 2 \\
-V_{0} \text { if } a / 2 \leq R_{s}<a
\end{array}\right.
$$

The on-axis version of the Rayleigh integral given by Eq. (8.3.1) is valid for any axisymmetric surface velocity. The discontinuous velocity distribution is handled by decomposing the integral into two parts,

$$
P_{\mathrm{axial}}(z)=i \omega \rho_{0} V_{0}\left[\int_{0}^{a / 2} \frac{e^{-i k\left(z^{2}+R_{s}^{2}\right)^{1 / 2}}}{\left(z^{2}+R_{s}^{2}\right)^{1 / 2}} R_{s} d R_{s}-\int_{a / 2}^{a} \frac{e^{-i k\left(z^{2}+R_{s}^{2}\right)^{1 / 2}}}{\left(z^{2}+R_{s}^{2}\right)^{1 / 2}} R_{s} d R_{s}\right]
$$

The change of variables $r^{2}=R_{s}^{2}+z^{2}$ is appropriate here, so we find that

$$
\begin{aligned}
P_{\text {axial }}(z) & =i \omega \rho_{0} V_{0}\left(\int_{z}^{\left(z^{2}+a^{2} / 4\right)^{1 / 2}} e^{-i k r} d r-\int_{\left(z^{2}+a^{2} / 4\right)^{1 / 2}}^{\left(z^{2}+a^{2}\right)^{1 / 2}} e^{-i k r} d r\right) \\
& =\rho_{0} c V_{0}\left(e^{-i k z}-2 e^{-i k\left(z^{2}+a^{2} / 4\right)^{1 / 2}}+e^{-i k\left(z^{2}+a^{2}\right)^{1 / 2}}\right)
\end{aligned}
$$

This expression resembles Eq. (8.3.6), but here there are three edge waves. The first term is a signal that departs from the center, the second arrives from the middle radius, and the third comes from the perimeter. The strength of each signal is proportional to the velocity discontinuity at the location where it departs. That is, at the center the surface velocity is $+V_{0}$, then it decreases by $-2 V_{0}$ at $R_{s}=a / 2$, and then it increases by $+V_{s}$ to return to the zero value on the baffle. This expression is not amenable to the type of manipulation that led to an explicit expression for the magnitude of the pressure, so we will merely evaluate it as a complex function.

Each complex exponential in the preceding has a different argument. It follows that the maxima and zeros of each term occur independently. Thus, at some locations they might all have maxima, leading to a pressure amplitude, that is, close to $4 \rho_{0} c\left|V_{0}\right|$, which is twice the maximum for a piston. It might also happen that at some locations the first and last term are maxima, while the second is a minimum, or vice versa, in which case the pressure would be close to zero. Figure 1, for $k a=10$, confirms this expectation. The peak pressure occurs close to the node for the circular piston, and no antinode is evident. However, the farfield pressure is approximately half that for a piston.


Figure 1, $k a=10$.
High values of $k a$ lead to the pressure being very small at some $z$, but not zero, as shown in Fig. 2 for $k a=40$. The pressure amplitude is at or close to $4 \rho_{0} c V_{0}$ at several locations, but the number of such locations is fewer than for a piston. Only one minimum is close to zero, and the pressure at another minimum is close to the maximum for the piston. Overall, the "antipiston" generates a larger pressure in the nearfield in comparison to the piston, but the pressure is lower in the farfield.


Figure 2, $k a=40$.

Whether this is a useful conceptual design depends on the application. If the intent is to generate a higher pressure on-axis in the nearfield that varies less and is larger overall than a piston, then the antipiston might be useful. In contrast, if one requires a large pressure in the farfield, then the antipiston is not a good idea.

### 8.4 An Overall Picture of the Pressure Field

At this juncture, we can evaluate the pressure at any point in the farfield. In contrast, the only aspect of the nearfield that we can evaluate readily is the pressure for a field point on-axis in the case where the surface vibration occurs over a circular region. An accurate determination of the pressure at an arbitrary location in the nearfield requires numerical methods. We shall do so here for the case of a circular piston, but the modifications to treat other systems are straightforward.

To integrate over $\mathcal{S}_{v}$ as required for the Rayleigh integral, Eq.(8.1.7), we will approximate the region as a set of patches. We will use the radial distance from the center of a patch to the field point to describe the contribution of that patch to the Rayleigh integral. There are more sophisticated numerical algorithms that use multiple points on a patch, analogously to Simpson's and higher order rules for integration over a single variable. The scheme we shall implement is not too difficult to program, but it will require a very fine grid. The same procedure was used by Zemanek ${ }^{6}$ to perform a detailed study that considered a wide range of parameters.

The procedure begins by dividing the circular area $\mathcal{S}_{v}$ into $N_{R}$ rings and a center circle, as shown in Fig. 8.10. The circle's radius is set to half the width $\Delta R$ of a ring, which leads to $\Delta R=2 a /\left(2 N_{R}+1\right)$.

Fig. 8.10 Decomposition of a circular region into patches


Each ring is divided into $N_{\theta}$ sectors, which leads to the sectorial increment being $\Delta \theta=2 \pi / N_{\theta}$. Each patch is associated with a pair of indices $j$ and $n$, such that its central point has polar coordinates $\left(R_{j}, \theta_{n}\right)$, where $R_{j}=n \Delta R$ and $\theta_{n}=(n-1 / 2) \Delta \theta$. The center circle is assigned to $j=n=0$, while $1 \leq j \leq N_{R}$ and $1 \leq n \leq N_{\theta}$ for the other patches. The inner and outer radii of ring $j$ are $r_{\text {in }}=(n-1 / 2) \Delta R$ and $r_{\text {out }}=(n+1 / 2) \Delta R$, and the angle subtended by a patch is $\Delta \theta$. Thus, the area of a patch is $\Delta \mathcal{S}_{j, n}=\left(r_{\text {out }}^{2}-r_{\text {in }}^{2}\right) \Delta \theta / 2=n \Delta \theta(\Delta R)^{2}$, with $\Delta \mathcal{S}_{0,0}=\pi(\Delta R)^{2} / 4$.

A map of the pressure in the nearfield is best obtained by depicting it in a cylindrical coordinate grid. The distance from a field point $\bar{x}_{0}$, whose cylindrical coordinates are ( $R, \theta, z$ ), to the central point $\bar{x}_{j, n}$ on a surface patch is

[^25]\[

$$
\begin{align*}
\left|\bar{x}_{0}-\bar{x}_{j, n}\right| & =\left[\left(R \cos \theta-R_{j} \cos \left(\theta_{n}\right)\right)^{2}+\left(R \sin \theta-R_{j} \sin \left(\theta_{n}\right)\right)^{2}+z^{2}\right]^{1 / 2} \\
& \equiv\left[R^{2}+R_{j}^{2}-2 R R_{j} \cos \left(\theta-\theta_{n}\right)+z^{2}\right]^{1 / 2} \tag{8.4.1}
\end{align*}
$$
\]

The integral in Eq.(8.1.7) is approximated as a sum of the contribution of each surface patch, with that contribution based on using $\left|\bar{x}-\bar{x}_{j, n}\right|$ for the entire patch. The resulting discretized form of the Rayleigh integral is

$$
\begin{align*}
P\left(\bar{x}_{0}\right)= & \frac{i \omega \rho_{0}}{2 \pi}\left\{V_{s}(0,0) \frac{e^{-i k\left|\bar{x}_{0}\right|}}{\left|\bar{x}_{0}\right|} \frac{\pi}{4}\left(\Delta R^{2}\right)+\right. \\
& \left.+\sum_{n=1}^{N_{\theta}} \sum_{j=1}^{N_{R}} V_{s}\left(R_{j}, \theta_{n}\right) \frac{e^{-i k\left|\bar{x}_{0}-\bar{x}_{j, n}\right|}}{\left|\bar{x}_{0}-\bar{x}_{j, n}\right|} n \Delta \theta(\Delta R)^{2}\right\} \tag{8.4.2}
\end{align*}
$$

Equation (8.4.2) may be used to evaluate the field resulting from any surface velocity distribution. In the special case of a translating piston, we set $V_{s}\left(R_{j}, \theta_{n}\right)=V_{0}$. The field in that case is axisymmetric, so it is sufficient to compute $P$ only on the $x z$ plane, where $\theta=0$. Another saving is that the $x z$ plane divides the surface into two parts whose contributions are equal, so only noncentral surface patches covering $0 \leq \theta \leq \pi$ need to be evaluated, with the result doubled. The field is computed on a grid of $R$ and $z$ values, and depicted as a three-dimensional surface plot, in which pressure is the third coordinate. The results of such computations are shown in Figs. 8.11 and 8.12. These computations used $N_{R}=400$ and $N_{\theta}=180$, that is, the sectors subtended $2^{\circ}$. These values were selected by progressively increasing each number until the plot of $|P|$ as a function of $z$ along $R=0$ matched the on-axis solution, Eq. (8.3.6). Equally important, the grid of field points must be sufficiently fine to capture the spatial fluctuations. The plotted results depict a grid whose points are separated by $\Delta z=a / 50$ in both directions. Another consideration is that placement of the field point on the vibrating surface will lead to a singularity in the integrand if the field point should happen to fall on one of the central points for a patch. This singularity can be addressed mathematically, as was done to derive the Surface Helmholtz Integral Equation. Here, the issue was avoided by $\operatorname{setting} \min (z)=\Delta z$. All of these considerations mean that the second plot, for $k a=24$, for which the maximum value of $z$ is $R_{0} / 2$, constitutes a grid of $6\left(10^{4}\right)$ field points, and the evaluation of the pressure at each field point required evaluation of a sum of $3.6\left(10^{4}\right)$ terms. Remarkably, this computation required less than four minutes on a desktop computer with dual quad 2.7 GHz Intel processors using MATLAB as the programming language.

The first case to be considered is $k a=6$, which is chosen because it is slightly lower than the frequency at which the on-axis pressure has a null. The maximum axial distance in Fig. 8.11 is twice the Rayleigh distance, and the width of the plotted region is twice the piston radius. The most evident feature is the smoothness of the spatial variation away from the piston's face. The farfield directivity, which is shown in Fig. 8.6, is discernible here too. For example, the pressure amplitude along $R=2 a$ is close to zero at $z \approx 2.4 a$, which corresponds to a polar angle of
$\tan ^{-1}(2 / 2.4)=39.8^{\circ}$, whereas the zero of $\mathcal{D}(\psi)$ at this frequency is $39.7^{\circ}$. Another feature worth noting is that the pressure very close to the baffle falls off quickly for transverse distances that are greater than $a$.


Fig. 8.11 Pressure amplitude in the nearfield of a piston in an infinite baffle, $k a=6$
The nearfield at high frequencies exhibits significant small-scale fluctuations, as may be seen in Fig. 8.12. This attribute is a consequence of cancelation between the signals that arrive at a field point from each location on the piston. As the frequency increases, the wavelength of these signals decreases, thereby increasing the


Fig. 8.12 Pressure amplitude in the nearfield of a piston in an infinite baffle, $k a=24$
possibility of constructive and destructive interference. On-axis, such interference is perfect at the nodes and antinodes and its transitions to the farfield directivity with increasing radial distance.

The spatial profiles in Fig. 8.13 provide a supplementary view of the data. The pressure out to $z=a$ fluctuates in the transverse direction, but it is reminiscent of what happens if a square wave is synthesized from a Fourier series that has an insufficient number of terms. Furthermore, the phase along these lines changes little for $R<a$, which is the characteristic of a simple plane wave. As we saw in the case when $k a=6$, the pressure close to the baffle falls off beyond $R>a$, but the decay is much more rapid at $k a=24$. At the larger distances, the farfield behavior can be seen to take over beyond the farthest maximum of the on-axis pressure. This distance is approximately half the Rayleigh distance.


Fig. 8.13 Spatial profiles of the pressure radiated by a piston in a baffle. The plotted data describe variations parallel and transverse to the axis of symmetry, $k a=24$

These trends are magnified in the case where $k a=72$, whose profiles are depicted in Fig. 8.14. The pressure within $R<a$ along the transverse lines close to the piston and at $z=a$ varies little and has a nearly constant phase. The general features are like those for $k a=24$.

The fall-off of pressure on the baffle beyond the piston radius explains why we may use analytical results for an infinite baffle model to describe a finite baffle. When $a$ is used to define dimensionless distances, an acoustic wavelength is $2 \pi /(k a)$. Maps such as Figs. 8.11 and 8.12 indicate that the pressure field close to the baffle, $z / a \ll 1$, is very small for $R / a>1+\pi /(k a)$. Furthermore, in this region, the pressure changes quite gradually with increasing $R$ and $z$. As a consequence, both particle velocity components are small in this region. Because the surface velocity is inherently small at sufficiently large $R / a$, there is no need for an infinite baffle to attain that condition; surrounding the piston with a rigid circular ring whose width equals the acoustic wavelength should be adequate to mimic the infinite baffle condition. The qualitative picture of a sound beam emerges from these computations. It is especially


Fig. 8.14 Spatial profiles of the pressure radiated by a piston in a baffle. The plotted data describes variations parallel and transverse to the axis of symmetry, $k a=72$
suitable as a description of the radiated sound field for a nonacoustician, although the fact that it is only correct in a macroscopic sense often is forgotten by practicing individuals. Figure 8.15 conveys the salient aspects.

Fig. 8.15 Conceptual model of the sound beam radiated by a circular piston in a baffle


The nearfield is considered to exist in the cylindrical region extending out from the piston's face to the Rayleigh distance. Within this region, the field is represented as a planar wave. The signal beyond the Rayleigh distance is considered to be a spherical wave within a cone whose semi-vertex angle is $\psi_{0}$, with the origin of this wave being the center of the piston. The angle is found from Fig. 8.15 to be $\tan ^{-1}\left(a / R_{0}\right)$. The model is intended for situations, where $k a \gg 1$, in which case $a / R_{0}=2 /(k a)$ is small. This reduces the arctangent to its argument, which leads to

$$
\begin{equation*}
\psi_{0} \approx \frac{a}{R_{0}}=\frac{2}{k a} \tag{8.4.3}
\end{equation*}
$$

For comparison, the maximum polar angle at which the intensity is at least half the on-axis maximum is half of the beamwidth, which was discussed in in Sect.8.2.2.

For large $k a$, this angle was found to be $\Delta_{\mathrm{bw}} / 2 \approx 1.62 / k a$. The conceptual angle $\psi_{0}$ for the radiated field is slightly greater than this angle.

EXAMPLE 8.4 Equation (8.2.30) gives a velocity distribution on the plane of a rigid baffle that supposedly describes the behavior of a transducer whose moving face translates but is not flush with the baffle. The function $Z(R)$ is the depth of the transducer's face at transverse distance $R$, with positive $Z(R)$ corresponding to the point being below the $x y$ plane. The discussion of Eq. (8.2.30) suggested that the phase delay associated with propagation from the transducer's face to the $x y$ plane may be ignored if $k a$ is sufficiently small. On the other hand, if $k a$ is very large, then the nearfield should be reasonably close to a plane wave, so the phase-shifted velocity distribution should be accurate. If so, then using it as the input to the Rayleigh integral should yield a reasonably accurate description of the pressure field. This hypothesis may be tested by considering the field of a hemispherical transducer mounted on a rigid baffle that executes a vibration as a rigid body. The radiated field was analyzed in Example 7.4 by formulating a spherical harmonic solution. The hemisphere's shape is $Z=-\left(a^{2}-R^{2}\right)^{1 / 2}$, where the minus sign applies because the hemisphere is above the $x y$ plane. Correspondingly, the approximate surface velocity on the baffle is

$$
V_{s}=\left\{\begin{array}{l}
V_{z} e^{i k\left(a^{2}-R^{2}\right)^{1 / 2}} \text { if } R<a \\
0 \text { if } R>a
\end{array}\right.
$$

This function could be used to map the entire field numerically in conjunction with Eq. (8.4.2). It is much simpler, although less definitive, if only the farfield and on-axis fields are evaluated according to the Rayleigh integral. Thus, the task is to evaluate the farfield and on-axis pressure by using the preceding expression for $V_{s}$ as the input to the Rayleigh integral, then to compare the results to those obtained from the spherical harmonic solution. Cases to consider are $k a=3$ and $k a=24$.

## Significance

Many applications use a transducer whose vibrating face does not lie in the plane of the baffle. Accurate evaluations of their pressure field requires numerical modeling techniques like those in Sect. 7.5. A simpler description like the one examined here would be especially useful for design studies.

## Solution

The farfield approximation of the Rayleigh integral for an arbitrary circular transducer is Eq. (8.2.22). The angular spectrum corresponding to the phase-shifted surface velocity is

$$
\tilde{f}(k \sin \psi)=\int_{0}^{a} e^{i k\left(a^{2}-R_{s}^{2}\right)^{1 / 2}} J_{0}\left(k R_{s} \sin \psi\right) R_{s} d R_{s}
$$

This expression cannot be integrated analytically, so we will perform a numerical integration based on the nondimensional distance $\xi=R_{s} / a$, which leads to

$$
\tilde{f}(k \sin \psi)=a^{2} \int_{0}^{1} e^{i k a\left(1-\xi^{2}\right)^{1 / 2}} J_{0}(k a \xi \sin \psi) \xi d \xi
$$

How to evaluate the integral depends on the software package to be used. The procedure implemented in MATLAB began by setting $k a$ and $\psi$. Then an anonymous function that evaluates the integrand was used as the input to another function that performs the integration. The anonymous function specification is $F=@$ ( $x i$ ) $\exp \left(1 i * k a * \operatorname{sqrt}\left(1-x i .{ }^{\wedge} 2\right)\right) . * b e s s e l j(0, k a * x i * \sin (p s i)) . * x i ;$ (The period before the exponent and the multiplication signs causes the operations to apply element by element to data vectors.) Then the integral is found according to quadl ( $\mathrm{F}, 0,1$ ). Both steps must be repeated for each value of $\psi$ from zero to $\pi / 2$ in order to construct a directivity pattern.

The on-axis version of the Rayleigh integral is Eq.(8.3.1). For the present $V_{s}$ function, this becomes

$$
P_{\text {axial }}(z)=i \omega \rho_{0} V_{z} \int_{0}^{a} e^{i k\left(a^{2}-R_{s}^{2}\right)^{1 / 2}} \frac{e^{-i k\left(z^{2}+R_{s}^{2}\right)^{1 / 2}}}{\left(z^{2}+R_{s}^{2}\right)^{1 / 2}} R_{s} d R_{s}
$$

Changing the integration variable to the radial distance for a field point slightly simplifies the integral. Thus, we define $a \xi=\left(z^{2}+R_{s}^{2}\right)^{1 / 2}$. This expression is solved for $R_{s}^{2}$ in order to transform the integral to

$$
P_{\text {axial }}(z)=i \omega \rho_{0} V_{z} \int_{z / a}^{\left(z^{2} / a^{2}+1\right)^{1 / 2}} e^{-i k a\left[\xi-\left(z^{2} / a^{2}+1-\xi^{2}\right)^{1 / 2}\right]} d \xi
$$

This too is not integrable analytically. A MATLAB anonymous function of $\xi$ that evaluates the integrand would be $\mathrm{G}=@$ (xi) $\exp (-1 i * k a *$ (xi-sqrt (z_a.^2+1xi.^2);. The integral correspondingly is evaluated according to quadl ( $G, z \_a$, sqrt $\left(z \_a .^{\wedge} 2+1\right)$ ), where $z \_a$ is the value of $z / a$ at the field point.

The spherical harmonic analysis of the pressure field radiated by a translating hemisphere in a rigid baffle was described in Example 7.4. The first operation evaluated the velocity coefficients $\left(V_{2 m}\right)_{\text {rigid }}$. Division of these values by the respective values of $H_{2 m}^{\prime}(k a)$ gave a set of pressure coefficients,

$$
B_{m}=\frac{\left(V_{2 m}\right)_{\mathrm{rigid}}}{i H_{2 m}^{\prime}(k a)}
$$

The resulting farfield directivity was

$$
\mathcal{D}_{\text {rigid }}(\psi)=\left|\frac{P}{\rho_{0} c V_{z}}\right|\left(\frac{r}{a}\right)=\frac{1}{k a}\left|\sum_{m=0}^{M / 2} B_{m} e^{i(2 m+1) \pi / 2} P_{2 m}(\cos \psi)\right|
$$

where $M$ is the minimum spherical harmonic index required for convergence to the desired precision. This series must be recomputed for each value of $\psi$.

The spherical harmonic series for the on-axis pressure is obtained by evaluating the general solution in Eq. (7.2.28) at $\psi=0$. The Legendre polynomials appearing in the series at this angle are $P_{2 m}(0) \equiv 1$, and $r=z$ on the axis of symmetry. Thus, the spherical harmonic series reduces to

$$
P_{\mathrm{axial}}(z)=\rho_{0} c v_{z} \sum_{m=0}^{M / s} B_{m} h_{2 m}\left(k a\left(\frac{z}{a}\right)\right)
$$

The results of this computation for a range of $z$ values beginning at the hemisphere, $z / a=1$, are plotted as $(z / a)\left|P /\left(\rho_{0} c V_{z}\right)\right|$. Doing so compensates for spherical spreading, which makes it easier to compare data over a large range of $z$ values.

Figure 1 compares the farfield and axial dependencies obtained from the approximate Rayleigh integral and the spherical harmonic series for $k a=3$. The directivity patterns are similar for all $\psi$, with the approximate on-axis value being less than 2 dB higher than the correct one. (This error increases rapidly as $k a$ is decreased below three.) The on-axis data tends to a constant value at $z / a \approx 3$, which suggests that this is the distance at which the farfield begins. For comparison, the Rayleigh distance is $R_{0}=k a / 2=1.5$.


Figure $1, k a=2$.

Setting $k a=24$ leads to very close agreement in the farfield directivity out to $\psi=60^{\circ}$, as can be seen in Fig. 2. The Rayleigh distance for a piston at this frequency is $R_{0} a=k a / 2=12$, which is consistent with the spherical harmonic prediction of the value of $z / a$ at which $r|P|$ on axis is a constant value. This is contrasted by the approximate model, which indicates that a constant value of $r|P|$ is attained at $z / a \approx 20$.


Figure 2, $k a=24$.

The data for the axial pressure suggests that the approximate analysis is quite good for the farfield, but exhibits significant errors in the nearfield. It is reasonable to generalize the preceding observations. If one has a transducer whose vibrating face is not flush to the plane of the baffle, it might be acceptable to use the Rayleigh integral with a phase-shifted velocity distribution to evaluate the farfield at high frequencies. Furthermore, although it might not be very accurate for the farfield at low frequencies and for the nearfield at any frequency, predictions obtained from this model give a qualitative picture that resembles the actual field. Such simulations might represent a useful alternative to numerical modeling techniques, and they can also be used to verify results obtained by such techniques.

### 8.5 Radiation Impedance of a Circular Piston

An important aspect of a transducer is the power required to drive it at a specified vibration amplitude. Aside from dissipation within the transducer, this is the same as the radiated power. Up to now, we have evaluated the time-averaged radiated power by using the farfield properties. We could perform a similar analysis here. Instead, we will determine the properties by analyzing the power that flows out of the piston. The virtue of such analysis is that it will provide a description of the instantaneous power output of the piston, which is important for such tasks as designing the electronics to drive the piston. We begin with a determination of the resultant force exerted by the fluid on the piston. This will lead to the radiation impedance, which we previously used in Sect.3.5.1 to form the end correction at the opening of a waveguide. We will now see how that correction is derived.

The radiation impedance $Z_{\text {rad }}$ for a vibrating piston is defined as the complex amplitude of the average pressure on the piston's face divided by the complex amplitude of the velocity, $V_{0}$. The average pressure is the total force divided by the area, and the total force is obtained by integrating $P d \mathcal{S}$ over the face. Thus,

$$
\begin{equation*}
Z_{\mathrm{rad}}=\frac{F_{s}}{\pi a^{2} V_{0}}, \quad F_{s}=\iint_{\mathcal{S}_{v}} P\left(\bar{x}_{s}\right) d \mathcal{S} \tag{8.5.1}
\end{equation*}
$$

A slightly different interpretation of this definition, which is useful as an extension to all transducers, is that the radiation impedance is the resultant force divided by the volume velocity. If we use the Rayleigh integral to evaluate $P$ at a point $\bar{x}$ on the surface, then that evaluation itself is an integral over the surface,

$$
\begin{equation*}
P\left(\bar{x}_{s}\right)=\frac{i \omega \rho_{0}}{2 \pi} V_{0} \iint_{\mathcal{S}_{v}} \frac{e^{-i k\left|\bar{x}_{s}-\bar{x}_{s}^{\prime}\right|}}{\left|\bar{x}_{s}-\bar{x}_{s}^{\prime}\right|} d \mathcal{S}^{\prime} \tag{8.5.2}
\end{equation*}
$$

Because $\bar{x}_{s}$ is situated somewhere on $S_{v}$, the integrand contains a singularity at $\bar{x}_{s}=\bar{x}_{s}^{\prime}$.

Rayleigh ${ }^{7}$ figured out an ingenious way to circumvent this singularity. Consider a pair of surface points $\bar{x}_{s}$ and $\bar{x}_{s}^{\prime}$. Let us define $\tau\left(\bar{x}_{s}, \bar{x}_{s}^{\prime}\right)$ as the contribution to pressure at $\bar{x}_{s}$ due to a unit surface velocity for a differential patch at $\bar{x}_{s}^{\prime}$, that is,

$$
\begin{equation*}
\tau\left(\bar{x}_{s}, \bar{x}_{s}^{\prime}\right)=\left(\frac{i \omega \rho_{0}}{2 \pi}\right) \frac{e^{-i k\left|\bar{x}_{s}-\bar{x}_{s}^{\prime}\right|}}{\left|\bar{x}_{s}-\bar{x}_{s}^{\prime}\right|} \tag{8.5.3}
\end{equation*}
$$

In terms of this quantity, $p\left(\bar{x}_{s}\right)$ is an integral $\tau\left(\bar{x}_{s}, \bar{x}_{s}^{\prime}\right) V_{0} d \mathcal{S}^{\prime}$ over $\mathcal{S}_{v}$, and the resultant force is an integral of $p\left(\bar{x}_{s}\right) d \mathcal{S}$, also over $\mathcal{S}_{v}$. Thus, $F_{s}$ may be found by evaluating

$$
\begin{equation*}
F_{s}=\iint_{\mathcal{S}_{v}} \iint_{\mathcal{S}_{v}} \tau\left(\bar{x}_{s}, \bar{x}_{s}^{\prime}\right) V_{0} d \mathcal{S}^{\prime} d \mathcal{S} \tag{8.5.4}
\end{equation*}
$$

Rayleigh's idea was to exploit the reciprocity property that $\tau\left(\bar{x}_{s}, \bar{x}_{s}^{\prime}\right)=\tau\left(\bar{x}_{s}, \bar{x}_{s}^{\prime}\right)$, which is obvious from its definition, as well as the principle of reciprocity. Rather than performing the double integral in a single operation, he decomposed it into two parts. To assist the discussion, we shall refer to $x_{s}^{\prime}$ as the source location and $\bar{x}_{s}$ as the pressure location. The first integral only considers sources that are closer to the center than the pressure location, that is, $\left|\bar{x}_{s}^{\prime}\right|<\left|\bar{x}_{s}\right|$. This is the attribute of the lightly shaded circle $\mathcal{S}_{\text {in }}$ in Fig. 8.16.

The second integral accounts for sources that are more distant than the pressure location. For every pair of points in the first integral, the same pair of points arise in the second integral. To recognize this, let $\bar{x}_{a}$ and $\bar{x}_{b}$ denote a specific pair of points, with $\bar{x}_{b}$ being the point that is closer to the center. Then, the first integral contains $\tau\left(\bar{x}_{a}, \bar{x}_{b}\right) V_{0} d \mathcal{S}_{b} d \mathcal{S}_{a}$, while the second integral contains $\tau\left(\bar{x}_{b}, \bar{x}_{a}\right) V_{0} d \mathcal{S}_{a} d \mathcal{S}_{b}$. Reciprocity of $\tau\left(\bar{x}_{b}, \bar{x}_{a}\right)$ tells us that these contributions are equal. A similar equality applies for every pair of points. Therefore, the evaluation of $F_{s}$ becomes one of the integrating over the circle $\mathcal{S}_{\text {in }}$ whose radius is the radial distance to an arbitrary point

[^26]Fig. 8.16 Decomposition of the face area of a vibrating piston for the purpose of evaluating the resultant force

$\bar{x}_{s}$ on its perimeter. The location of this perimeter point is held fixed in the integral over $\mathcal{S}_{\text {in }}$. It follows that $\bar{x}_{s}^{\prime}$ is the only position variable on which this integral depends. An integral of this functional dependence over $\mathcal{S}_{\text {in }}$ accounts for all sources that are closer than the pressure location. Doubling that result accounts for all source locations outside the pressure location, so the evaluation becomes

$$
\begin{equation*}
F_{s}=2 \iint_{\mathcal{S}_{v}} \iint_{\mathcal{S}_{\text {in }}} \tau\left(\bar{x}_{s}, \bar{x}_{s}^{\prime}\right) d \mathcal{S}^{\prime} d \mathcal{S} \tag{8.5.5}
\end{equation*}
$$

The second part of Rayleigh's contribution is the way in which the integral over $\mathcal{S}_{\text {in }}$ is carried out. The definition of $\tau\left(\bar{x}, \bar{x}_{s}^{\prime}\right)$ depends on the distance between the points. Recall that $\bar{x}_{s}$ is held fixed in the integral over $S_{\text {in }}$. Designating it as the origin for a set of polar coordinates $\left(R_{s}, \phi\right)$ is very useful because $R_{s}=\left|\bar{x}_{s}-\bar{x}_{s}^{\prime}\right|$, as depicted in Fig. 8.16. The integration is performed first over $R_{s}$, then over $\phi$. The length of diameter $O A$ is $2 r \equiv 2\left|\bar{x}_{s}\right|$. This diameter is the hypotenuse of right triangle $O A B$, so the range of radial distances is $0 \leq R_{s} \leq 2 a \cos \phi$, and the range of polar angles is $-\pi / 2 \leq \phi \leq \pi / 2$. In this polar coordinate system, $d \mathcal{S}^{\prime}=R_{s} d R_{s} d \phi$, which cancels the occurrence of $R_{s}$ in the denominator of $\tau\left(\bar{x}, \bar{x}_{s}^{\prime}\right)$, thereby removing the singularity. Thus, we find that

$$
\begin{align*}
\iint_{\mathcal{S}_{\text {in }}} \tau\left(\bar{x}_{s}, \bar{x}_{s}^{\prime}\right) d \mathcal{S}^{\prime} & =\frac{i \omega \rho_{0}}{2 \pi} V_{0} \int_{-\pi / 2}^{\pi / 2} \int_{0}^{2 r \cos \phi} e^{-i k R_{s}} d R_{s} d \phi \\
& =\frac{1}{2 \pi} \rho_{0} c V_{0} \int_{-\pi / 2}^{\pi / 2}[1-\cos (2 k r \cos \phi)+i \sin (2 k r \cos \phi)] d \phi \tag{8.5.6}
\end{align*}
$$

Because $\cos \phi$ is an even function relative to $\phi=0$, the integral is twice the integral over $0 \leq \phi \leq \pi / 2$, so that

$$
\begin{align*}
\iint_{\mathcal{S}_{\text {in }}} \tau\left(\bar{x}_{s}, \bar{x}_{s}^{\prime}\right) d \mathcal{S}^{\prime} & =\frac{1}{\pi} \rho_{0} c V_{0} \int_{0}^{\pi / 2}[1-\cos (2 k r \cos \phi)+i \sin (2 k r \cos \phi)] d \phi \\
& =\frac{1}{\pi} \rho_{0} c V_{0}\left[\frac{\pi}{2}-\mathcal{I}_{c}+i \mathcal{I}_{s}\right] \\
\mathcal{I}_{c}=\int_{0}^{\pi / 2} & \cos (2 k r \cos \phi) d \phi, \quad \mathcal{I}_{s}=\int_{0}^{\pi / 2} \sin (2 k r \cos \phi) d \phi \tag{8.5.7}
\end{align*}
$$

An integral like $\mathcal{I}_{c}$ arose in the derivation of the farfield directivity of a ring. The tabulated integral ${ }^{8}$ used there extended over $0 \leq \phi \leq \pi$. The integrand of $\mathcal{I}_{c}$ is an even function with respect to $\phi=\pi / 2$, so $\mathcal{I}_{c}$ is half the value of the tabulated entry, that is,

$$
\begin{equation*}
\mathcal{I}_{c}=\frac{\pi}{2} J_{0}(2 k r) \tag{8.5.8}
\end{equation*}
$$

The integration range for $\mathcal{I}_{s}$ does not match any entry for Bessel functions in a basic resource for special functions. However, it does appear in a chapter on Struve functions, which are denoted as $\mathbf{H}_{\nu}(x)$, where $\nu$ is the order. Alternative definitions of this function describe it as a power series or as an integral, ${ }^{9}$

$$
\begin{align*}
\mathbf{H}_{\nu}(x) & =\left(\frac{x}{2}\right)^{\nu+1} \sum_{k=0}^{\infty} \frac{\left(-x^{2} / 4\right)^{k}}{\Gamma(k+3 / 2)(k+\nu+3 / 2)}  \tag{8.5.9}\\
& =\frac{2(x / 2)^{\nu}}{\pi^{1 / 2} \Gamma(\nu+1 / 2)} \int_{0}^{\pi / 2} \sin (x \cos \theta)(\sin \theta)^{2 \nu} d \theta
\end{align*}
$$

These alternative definitions feature the Gamma function, denoted as $\Gamma(z)$. It is another special function described in our basic reference. ${ }^{10}$ It follows the recurrence relation $\Gamma(z+1)=z \Gamma(z)$. Thus, for integer arguments, it is related to a factorial according to $\Gamma(n+1)=n$ ! For a noninteger positive argument $z$, it may be computed by evaluating the function at the fractional part of the argument, $z_{f}=z-\operatorname{floor}(z)$, followed by application of the recurrence relation. The fractional part for the preceding formulas is $z_{f}=1 / 2$ and $\Gamma(1 / 2)=\pi^{1 / 2}$. If the chosen mathematical software does not provide a routine that evaluates the Struve function, it may be evaluated numerically according to either representation in Eq. (8.5.9). However, both forms will encounter difficulty with round-off error if the argument is large, e.g. $x>40$. An asymptotic approximation is quite effective in such situations.

A comparison of the integral representation in Eq. (8.5.9) and the definition of $\mathcal{I}_{s}$ show that

$$
\begin{equation*}
\mathcal{I}_{s}=\frac{\pi^{1 / 2} \Gamma(1 / 2)}{2} \mathbf{H}_{0}(2 k r) \equiv=\frac{\pi}{2} \mathbf{H}_{0}(2 k r) \tag{8.5.10}
\end{equation*}
$$

[^27]Correspondingly, we find that

$$
\begin{equation*}
\iint_{\mathcal{S}_{\text {in }}} \tau\left(\bar{x}_{s}, \bar{x}_{s}^{\prime}\right) V_{0} d \mathcal{S}^{\prime}=\frac{1}{2} \rho_{0} c V_{0}\left[1-J_{0}(2 k r)+i \mathbf{H}_{0}(2 k r)\right] \tag{8.5.11}
\end{equation*}
$$

The total resultant force is found by integrating this term over $\mathcal{S}_{v}$. Because $r$ is the distance from the center to $\bar{x}_{s}$, the integration is done by changing to a polar coordinate system centered on $\mathcal{S}_{v}$. Thus, we must evaluate

$$
\begin{equation*}
F_{s}=\iint_{\mathcal{S}_{v}} 2 \iint_{\mathcal{S}_{\text {in }}} \tau\left(\bar{x}_{s}, \bar{x}_{s}^{\prime}\right) d \mathcal{S}^{\prime} d \mathcal{S}=\int_{0}^{a} \int_{-\pi}^{\pi} \rho_{0} c V_{0}\left[1-J_{0}(2 k r)+i \mathbf{H}_{0}(2 k r)\right] r d \theta d r \tag{8.5.12}
\end{equation*}
$$

The $\theta$ integration gives a $2 \pi$ factor. To integrate the Bessel and Struve function terms, we invoke a recurrence relation that has a similar form for both, specifically,

$$
\begin{equation*}
\frac{d}{d y}\left(y^{n} J_{n}(y)\right)=y^{n} J_{n-1}(y), \quad \frac{d}{d y}\left(y^{n} \mathbf{H}_{n}(y)\right)=y^{n} \mathbf{H}_{n-1}(y) \tag{8.5.13}
\end{equation*}
$$

The result is

$$
\begin{equation*}
F_{s}=\pi a^{2} V_{0} Z_{\mathrm{rad}}, \quad Z_{\mathrm{rad}}=\rho_{0} c \chi(k a) \tag{8.5.14}
\end{equation*}
$$

where the specific impedance is

$$
\begin{equation*}
\chi(k a)=\left(1-\frac{J_{1}(2 k a)}{k a}\right)+i\left(\frac{\mathbf{H}_{1}(2 k a)}{k a}\right) \tag{8.5.15}
\end{equation*}
$$

Figure 8.17 repeats from Chap. 3 the graph of $\chi$ as a function of $k a$.


Fig. 8.17 Plot of the frequency dependence of the specific acoustic impedance $\chi$ for a vibrating circular piston of radius $a$

For small $k a$, the first few terms of a Taylor series expansion are adequate,

$$
\begin{equation*}
\chi(k a)=\left[\frac{1}{2}(k a)^{2}-\frac{1}{12}(k a)^{4}+\cdots\right]+i \frac{4}{\pi}\left[\frac{2}{3}(k a)-\frac{8}{45}(k a)^{3}+\cdots\right] \tag{8.5.16}
\end{equation*}
$$

The specific impedance tends to $\chi(k a)=1$ as $k a \rightarrow \infty$, which is merely the plane wave limit associated with high frequencies.

One reason that the specific impedance is important is that it is needed to predict the voltage and current required to drive the piston transducer at a specified velocity. Knowledge of the specific impedance provides a simple way to evaluate the power required to drive a piston. Aside from losses within the transducer, this is the same as the power radiated by the piston. Let us consider this quantity as a function of time. All points on the piston's face translate at the same velocity, so the instantaneous power is the product of the resultant force and that velocity. The instantaneous power input to the fluid at the piston face, therefore, is

$$
\begin{align*}
\mathcal{P}(t) & =\frac{1}{4} \operatorname{Re}\left[\left(F_{s} e^{i \omega t}+\text { c.c. }\right)\left(V_{0} e^{i \omega t}+\text { c.c. }\right)\right] \\
& =\frac{1}{2} \operatorname{Re}\left[F_{s}\left(V_{0}\right)^{*}+F_{s} V_{0} e^{2 i \omega t}\right]  \tag{8.5.17}\\
& =\frac{\pi a^{2}}{2}\left[\left(\operatorname{Re} Z_{\mathrm{rad}}\right)\left|V_{0}\right|^{2}+\operatorname{Re}\left(\left(V_{0}\right)^{2} Z_{\mathrm{rad}} e^{2 i \omega t}\right)\right]
\end{align*}
$$

The invariant part of this expression is the time-averaged power,

$$
\begin{equation*}
\mathcal{P}_{\mathrm{av}}=\frac{\pi a^{2}}{2} \operatorname{Re}\left(Z_{\mathrm{rad}}\right)\left|V_{0}\right|^{2} \tag{8.5.18}
\end{equation*}
$$

The fluctuating part is the power that flows into and out of the fluid across the piston's face. It oscillates at a frequency $2 \omega$ with an amplitude that is $\left(\pi a^{2} / 2\right)\left|Z_{\mathrm{rad}}\right|\left|V_{0}\right|^{2}$, so

$$
\begin{equation*}
\left|\mathcal{P}_{\text {fluct }}\right|=\mathcal{P}_{\mathrm{av}} \frac{\left|Z_{\mathrm{rad}}\right|}{\operatorname{Re}\left(Z_{\mathrm{rad}}\right)} \tag{8.5.19}
\end{equation*}
$$

It follows that the average power flow across the piston face never exceeds the fluctuating part, so there is a time interval during which the power flows from the fluid into the piston. In this condition, the pressure resultant acting on the piston is in the same sense as the piston's velocity.

It is enlightening to compare the power flow across the piston face to the power that flows across a very large surrounding hemisphere. We know that the average portions should be the same, but what about the fluctuating parts? To answer this, we us a large hemisphere whose radius is sufficiently large that we may use the farfield approximation that $P / V_{r}=\rho_{0} c$ at any polar angle. The time-domain version of Eq. (8.2.23) is

$$
\begin{equation*}
p_{\mathrm{ff}}\left(\bar{x}_{0}, t\right)=\rho_{0} c \frac{J_{1}(k a \sin \psi)}{\sin \psi} \frac{a}{r}\left(\frac{1}{2}\right)\left(V_{0} e^{-i k r} e^{i \omega t}+\text { c.c. }\right) \tag{8.5.20}
\end{equation*}
$$

The instantaneous radial intensity is $p_{\mathrm{ff}}^{2} /\left(\rho_{0} c\right)$, and the field is axisymmetric, so $d \mathcal{S}_{0}=2 \pi r^{2} d \psi$. Hence, the instantaneous power flowing across a hemisphere of radius $r$ is given by

$$
\begin{align*}
\mathcal{P}_{\mathrm{ff}}(t) & =\int_{0}^{\pi} \frac{1}{\rho_{0} c}\left[p_{\mathrm{ff}}\left(\bar{x}_{0}\right)\right]^{2} 2 \pi r^{2} d \psi \\
& =\pi \rho_{0} c a^{2}\left[\int_{0}^{\pi}\left(\frac{J_{1}(k a \sin \psi)}{\sin \psi}\right)^{2} d \psi\right]\left[\left|V_{0}\right|^{2}+\operatorname{Re}\left(\left(V_{0}\right)^{2} e^{2 i(k r-\omega t)}\right)\right] \tag{8.5.21}
\end{align*}
$$

There is no need to evaluate the integral because the time-averaged power is the constant part of this expression. The net power flowing into any region of an ideal fluid must have a zero average value, so this portion of $\mathcal{P}_{\mathrm{ff}}(t)$ must equal $\mathcal{P}_{\mathrm{av}}$ in Eq. (8.5.18). Hence, writing $V_{0}$ in polar form leads to

$$
\begin{equation*}
\mathcal{P}_{\mathrm{ff}}(t)=\frac{\pi a^{2}}{2} \operatorname{Re}\left(Z_{\mathrm{rad}}\right)\left|V_{0}\right|^{2}\left[1+e^{2 i\left(k r-\omega t+\arg \left(V_{0}\right)\right)}\right] \tag{8.5.22}
\end{equation*}
$$

This description of $\mathcal{P}_{\mathrm{ff}}(t)$ highlights the fact that the fluctuating part has an amplitude that equals the mean value. This is different from the behavior at the transducer face, where the magnitude of the fluctuating part of the power flow exceeds the average. At any instant, the difference between the fluctuating parts of the power radiated by the piston and the power that flows into the farfield represents the rate of change of the fluid's mechanical energy in the region close to the piston.

EXAMPLE 8.5 A circular 450 mm diameter piston is mounted in a baffle on the floor in an anechoic chamber. In one test, the electrical signal creates a square wave alternating between $+V_{0}$ and $-V_{0}$ with a fundamental frequency of 1 kHz . The sound pressure on-axis 4 meters from the piston is measured to be $125 \mathrm{~dB} / 20 \mu \mathrm{~Pa}$. This measurement is cutoff at 18 kHz in order to account for audible tones only. Determine the power output as a function of time. Also determine the waveform measured at the 4 m distance on-axis. Graph both time-dependent quantities.

## Significance

Beyond the obvious objective of performing an evaluation of radiated power, we will see that power requirements are significantly affected by the harmonic content. In addition, the pressure waveform provides a preview of an interesting phenomenon that arises in a time-domain analysis.

## Solution

The value of $V_{0}$ is not specified. We will determine it by evaluating the pressure at $z=4 \mathrm{~m}$ on-axis in terms of $V_{0}$, which will yield a value of the mean-squared pressure at that location. The square wave is specified to be a periodic signal defined by

$$
\begin{equation*}
v_{s}(t)=V_{0}[h(t)-2 h(t-T / 2)] \text { if } 0 \leq t \leq T, \quad v_{s}(t \pm T)=v_{s}(t) \tag{1}
\end{equation*}
$$

The value of $\omega_{1}$ is $2000 \pi \mathrm{rad} / \mathrm{s}$. The Fourier series representation of this signal is

$$
\begin{gather*}
v_{s}=\operatorname{Re} \sum_{n=1}^{\infty} V_{n} e^{i \omega_{n} t}, \quad \omega_{n}=n \omega_{1}, \quad \omega_{1}=\frac{2 \pi}{T}  \tag{2}\\
V_{n}=\frac{2}{i \pi n}\left[1-(-1)^{n}\right] V_{0}
\end{gather*}
$$

Only the $V_{n}$ values for odd $n$ are not zero. Each harmonic contained in the surface vibration is associated with a different wavenumber,

$$
\begin{equation*}
k_{n}=n \omega_{1} / c \tag{3}
\end{equation*}
$$

Each value of $k_{n}$ leads to a contribution to the on-axis pressure according to Eq. (8.3.6). The resulting Fourier time series is

$$
\begin{gather*}
p(z, t)=\operatorname{Re} \sum_{n \text { odd }}^{\infty} P_{n}(z) e^{i n \omega_{1} t} \\
P_{n}(z)=2 i \rho_{0} c\left(\frac{4 V_{0}}{i \pi n}\right) e^{-i\left(k_{n} / 2\right)\left[z+\left(z^{2}+a^{2}\right)^{1 / 2}\right]} \sin \left[\frac{k_{n}}{2}\left(\left(z^{2}+a^{2}\right)^{1 / 2}-z\right)\right] \tag{4}
\end{gather*}
$$

We take $c=340 \mathrm{~m} / \mathrm{s}$ and $\rho_{0}=1.2 \mathrm{~kg} / \mathrm{m}^{3}$ as representative values for air. The pressure coefficients at $z=4 \mathrm{~m}, R=0$ of the odd harmonics are tabulated below. They are cutoff at 18 kHz , which for a 1 kHz fundamental corresponds to $n=18$.

| $n$ | 1 | 3 | 5 | 7 | 9 |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\operatorname{Re}\left(\left.\frac{P_{n}}{V_{0}}\right\|_{z=4}\right)$ | 9.1155 | -26.4029 | 40.9712 | -51.3675 | 56.6292 |
| $\operatorname{Im}\left(\left.\frac{P_{n}}{V_{0}}\right\|_{z=4}\right)$ | 59.9788 | -54.3143 | 43.6173 | -29.0723 | 12.2696 |
| $\operatorname{Re}\left(\left.\frac{P_{n}}{V_{0}}\right\|_{z=4}\right)$ | -56.3883 | 50.9007 | -40.9956 | 27.9556 |  |
| $\operatorname{Im}\left(\left.\frac{P_{n}}{V_{0}}\right\|_{z=4}\right)$ | 4.9911 | -20.9190 | 33.9445 | -42.8971 |  |

The given sound pressure level of 125 dB at $z=4 \mathrm{~m}$ corresponds to a meansquared pressure of $1.2649\left(10^{3}\right) \mathrm{Pa}^{2}$. At the same time, the mean-squared pressure is the incoherent sum of the mean-squared pressure of each harmonic in the tabulation, that is,

$$
\left(p^{2}\right)_{\mathrm{av}}=\left.\frac{1}{2}\left|V_{0}\right|^{2} \sum_{n \text { odd }}^{18}\left|\frac{P_{n}}{V_{0}}\right|_{z=4}\right|^{2}=14719\left|V_{0}\right|^{2}=1.2649\left(10^{3}\right)
$$

From this, we find that the piston's velocity amplitude is

$$
V_{0}=0.2932 \mathrm{~m} / \mathrm{s}
$$

With this value determined, the actual $V_{n}$ and $P_{n}$ coefficients can be computed. The primary factor deciding the time increment for synthesis of the Fourier series is how well we wish to capture the rise and fall of the square wave. Let us select $\Delta t=T / 100$ for this purpose. Presumably, the full harmonic content of the pressure is measured and then filtered for the content below 18 kHz . It is possible that despite such filtering, the existence of higher harmonics is important, especially for considerations related to the power output of the transducer. To ascertain whether such speculation is correct, all series shall be synthesized for a series length of $N=100$, as well for $N=18$. The first evaluation tests the adequacy of either truncation by evaluating the input velocity waveform according to Eq. (2).

We could use an inverse FFT routine to create the waveforms, but we shall implement the matrix algorithm in Eq. (1.4.39). Thus, $v_{s}$ at instant $t_{j}$ is indicated by Eq. (1) to be

$$
\begin{equation*}
v_{s}\left(t_{j}\right)=\operatorname{Re} \sum_{n \text { odd }}^{N} V_{n} e^{i \omega_{n} t_{j}} \Longrightarrow\left\{v_{s}\right\}=[E]^{\mathrm{H}}\{V\} \tag{5}
\end{equation*}
$$

where

$$
\{V\}=\left[\begin{array}{lll}
V_{1} & V_{3} & \cdots
\end{array}\right]^{\mathrm{T}}, \quad[E]=\left[\begin{array}{ccc}
e^{-i \omega_{1} t_{1}} & e^{-i \omega_{1} t_{2}} & \cdots \\
e^{-i \omega_{2} t_{1}} & e^{-i \omega_{2} t_{2}} & \cdots \\
\vdots & \vdots & \ddots
\end{array}\right]
$$

The superscript in Eq. (5) specifies the Hermitian of $[E]$, which is the complex conjugate transpose of the array.

Figure 1 shows that either truncation gives an adequate reproduction of the square wave, although the ringing for the smaller $N$ is noticeable. Note that this ringing stems from the fact that $N=18$ does not meet the Nyquist criterion.


Figure 1.

Figure 2 depicts the pressure waveform at $z=4 \mathrm{~m}$. It was computed by adapting the matrix algorithm in Eq. (5) to Eq. (4). The result is somewhat surprising because it indicates that the waveform is a sequence of spikes. One clue why this is so is found in the Rayleigh distance $\left(R_{0}\right)_{n}=k_{n} a^{2} / 2$ for each harmonic. This distance for
the first harmonic is 0.468 m , and it is not until $n=9$ that $z=4 \mathrm{~m}$ is less than the Rayleigh distance. Thus, for the most significant harmonics, the field point is in the farfield. According to Eq. (8.3.11), the contribution of each harmonic at fixed $z$ in the farfield is proportional to $\left(R_{0}\right)_{n} V_{n}$. The $V_{n}$ values are inversely proportional to $n$, and the Rayleigh distance of each harmonic is proportional to $\omega_{n}$, which means that the lower $P_{n}(z)$ values should have nearly equal magnitude. This expectation is borne out by the tabulation of $P_{n}(z) / V_{0}$. A Fourier series whose coefficients have equal magnitude, with suitable relative phase delays, represents a periodic train of impulses. This is manifested as the pressure being a series of alternating spikes. But why does this happen? That explanation will be found in the next section, which explores the time-domain Rayleigh integral.


Figure 2.
It is not necessary to synthesize waveforms to evaluate the time-averaged power because the intensity is the incoherent sum of contributions from each harmonic. Therefore, $\mathcal{P}_{\text {av }}$ is the sum of terms described by Eq. (8.5.18), with $Z_{\mathrm{rad}}$ for each $n$ evaluated at the respective $k_{n}$ value. That is,

$$
\begin{equation*}
\mathcal{P}_{\mathrm{av}}=\frac{\pi a^{2} \rho_{0} c}{2} \operatorname{Re} \sum_{n \text { odd }}^{N} \chi\left(k_{n} a\right)\left|V_{n}\right|^{2} \tag{5}
\end{equation*}
$$

The result for $N=100$ is $\mathcal{P}_{\text {av }}=5.27 \mathrm{~W}$. Cutting off the series at $N=18$ yields $\mathcal{P}_{\mathrm{av}}=5.17 \mathrm{~W}$, so the shorter series length seems to be quite adequate.

Unlike the evaluation of the average power, interference between harmonics is an important feature of the instantaneous power. We begin this evaluation by describing the force resultant as a Fourier series. The complex amplitude for a harmonic is described by Eq. (8.5.14), which we adapt by evaluating $Z_{\mathrm{rad}}$ at the respective harmonic frequency and multiplying it by the velocity coefficient $V_{n}$ in Eq. (2). The series is

$$
\begin{equation*}
F_{s}=\pi a^{2} \rho_{0} c \operatorname{Re} \sum_{n \text { odd }}^{N} \chi\left(k_{n} a\right) V_{n} e^{i n \omega_{1} t} \tag{6}
\end{equation*}
$$

The matrix algorithm in Eq. (5) for this calculation is

$$
\begin{equation*}
\left\{F_{s}\right\}=[E]^{\mathrm{H}}\{\{\chi\} \cdot\{V\}\} \tag{8.5.22}
\end{equation*}
$$

where $\{\{\chi\} \cdot\{V\}\}$ denotes an element by element multiplication.

The instantaneous power may be computed by multiplying the $F_{s}$ value at any $t$ by the $v_{s}$ value plotted in Fig. 1. The result of both computations is shown in Fig. 3. The force waveform resembles the piston's velocity, except that the plateaus are not flat. The radiated power is positive if $v_{s}$ and $F_{s}$ have the same sign, so the plot of $\mathcal{P}_{\text {rad }}$ is a periodic sequence of positive plateaus, with $T / 2$ as the period. The instants at which $\mathcal{P}_{\text {rad }}$ is zero are the instants at which $v_{s}=0$. There is no interval in which $P_{\mathrm{rad}}$ is negative. This is so because all $k_{n} a>3$, so all $\chi\left(k_{n} a\right)$ are close to unity. The maximum radiated power, which occurs soon after the transitions of $v_{s}$, is 6.18 W . The power falls off after the peak, then drops off rapidly at the velocity transitions. The consequence is that the average power does not differ much from the peak value. An interesting aspect is that the inadequacies of using the smaller $N$ are evident for the force, and especially the power. The shorter series length leads to $\max \left(\mathcal{P}_{\text {rad }}\right)=8.84$ W , which is more than $40 \%$ high.


Figure 3.

### 8.6 Time Domain Rayleigh Integral

A time-domain description of the signal generated by a vibrating piston has advantages in some circumstances. One approach describes the field as a superposition of plane and evanescent waves of the type studied in Sect.5.1. A two-dimensional Fourier transform of the pressure leads to a representation that matches the angular spectrum of the surface velocity. The requirement that the pressure at each wavenumber matches the normal surface velocity yields a representation of the pressure as an inverse Fourier transform. From a philosophical viewpoint, this formulation is attractive because it provides a similar perspective to that provided by the prior investigations of radiation from spheres to cylinders. The formulation is primarily intended for numerical evaluations, for which FFT technology substantially reduces
the computational effort. The paper by Wu, Kazys, and Stepinski ${ }^{11}$ explains the details. We shall not delve into this line of investigation because it does not lead to new insights as to the nature of the pressure field.

In contrast, the time-domain version of the Rayleigh integral provides a different perspective. For example, it provides a simple explanation of the impulse-like waveform observed in the previous example. However, evaluation of the pressure field at an arbitrary field point is more difficult than it is in the frequency-domain approach.

We have regarded the Rayleigh integral as a prescription for evaluating the frequency-domain response when the surface vibration is a set of point sources associated with $\operatorname{Re}\left(V_{s}\left(\bar{x}_{s}\right) \exp (i \omega t)\right)$. A different perspective views the Rayleigh integral as the Fourier transform of a time-domain representation of the radiated field. Our task is to determine what that time representation is.

We shall not call on the inverse Fourier transform to carry out this determination. Rather, we will use the notion of function/transform pairs for operations in the time domain, such as those in Table B1 of Appendix B. The first step is to recognize that the factor $i \omega V_{s}\left(\bar{x}_{s}\right)$ in the Rayleigh integral is the Fourier transform of the rate of change of $v_{s}\left(\bar{x}_{s}, t\right)$, that is, $\mathcal{F}\left(\left(\partial v_{s} / \partial t\right), \omega\right)=i \omega V_{s}\left(\bar{x}_{s}\right)$. Hence, the Rayleigh integral may be written as

$$
\begin{equation*}
P\left(\bar{x}_{0}\right)=\frac{\rho_{0}}{2 \pi} \iint_{\mathcal{S}} \mathcal{F}\left(\frac{\partial}{\partial t} v_{s}\left(\bar{x}_{s}, t\right), \omega\right) \frac{e^{-i k\left|\bar{x}_{0}-\bar{x}_{s}\right|}}{\left|\bar{x}_{0}-\bar{x}_{s}\right|} d \mathcal{S} \tag{8.6.1}
\end{equation*}
$$

The other occurrence of $\omega$ is in the complex exponential. The coefficients of a Fourier series that are delayed by time $\tau$ relative to another function are described in Eq. (1.4.11). It is shown there that if $p(t)$ and $q(t)$ are two functions that have the same period $T$, with $q(t)=p(t-\tau)$, then their Fourier series coefficients are related by $Q_{j}=P_{j} e^{-i \omega_{j} \tau}$, where $\omega_{j}=j(2 \pi / T)$. A Fourier transform corresponds to the coefficients of the Fourier series of a function whose period is infinite, which replaces the discrete frequencies $\omega_{j}$ with a continuous variable $\omega$. Thus, if $g(t)=g(t-\tau)$, their Fourier transforms are related by $\mathcal{F}(g(t), \omega)=\mathcal{F}(f(t), \omega) e^{-i \omega \tau}$. (This property also may be identified from the coordinate shifting property in Table B1 of Appendix B, except that the sign of the exponent must be reversed because the tabulation is based on our definition of a spatial Fourier transform.) Because $k=\omega / c$, the exponential in the Rayleigh integral corresponds to delaying a function by $\left|\bar{x}_{0}-\bar{x}_{s}\right| / c$. Correspondingly, the preceding form of the Rayleigh integral may be written as

$$
\begin{equation*}
P\left(\bar{x}_{0}\right)=\frac{\rho_{0}}{2 \pi} \iint_{\mathcal{S}} \mathcal{F}\left(\frac{\partial}{\partial t} v_{s}\left(\bar{x}_{s}, t-\frac{\left|\bar{x}_{0}-\bar{x}_{s}\right|}{c}\right), \omega\right) \frac{1}{\left|\bar{x}_{0}-\bar{x}_{s}\right|} d \mathcal{S} \tag{8.6.2}
\end{equation*}
$$

[^28]By definition, the inverse Fourier transform of a Fourier transform is the original time-domain function. Because the source-field point distance is independent of $\omega$, it is constant in the transformation. We thereby obtain the time domain Rayleigh integral,

$$
\begin{equation*}
p\left(\bar{x}_{0}\right)=\frac{\rho_{0}}{2 \pi} \iint_{\mathcal{S}} \frac{1}{\left|\bar{x}_{0}-\bar{x}_{s}\right|} \frac{\partial}{\partial t} v_{s}\left(\bar{x}_{s}, t-\frac{\left|\bar{x}_{0}-\bar{x}_{s}\right|}{c}\right) d \mathcal{S} \tag{8.6.3}
\end{equation*}
$$

This form is readily understood in terms of source distributions. Consider a differential element of the surface. The mass acceleration of this element is $\rho_{0} d \mathcal{S} d v_{s} / d t$, and the distance from it to the field point is $r=\left|\bar{x}_{0}-\bar{x}_{s}\right|$. The time-domain pressure radiated by this element is double the value in Eq. (6.4.3) because the sound radiates into a half space, rather than a free space. Thus, it is

$$
\begin{equation*}
d p\left(\bar{x}_{0}, t\right)=\frac{\rho_{0}}{2 \pi r} \frac{\partial}{\partial t} d Q_{s}\left(t-\frac{r}{c}\right) \tag{8.6.4}
\end{equation*}
$$

The time-domain Rayleigh integral is the superposition of the signals radiated by a sheet of these sources distributed over the surface. The signals from different sources that arrive contemporaneously at a field point departed from the surface at different instants. Specifically, the signal from a differential source at $\bar{x}_{s}$ that arrives at the field point at time $t$ must have left the surface at time $t-\left|\bar{x}_{0}-\bar{x}_{s}\right| / c$. Hence, the signal observed at a field point at time $t$ consists of contributions that departed from the surface in a time interval beginning at $t-\max \left(\left|\bar{x}_{0}-\bar{x}_{s}\right|\right) / c$ and ending at $t-\min \left(\left|\bar{x}_{0}-\bar{x}_{s}\right|\right) / c$.

An evaluation of the Rayleigh integral for an arbitrary location and surface velocity requires numerical methods. The technique described by Fig. 8.10 also is suitable for this application. The radial distance from the central point on a surface patch to the selected field point is used to evaluate $t-r / c$ for that patch. Then, Eq.(8.6.4) is added for each patch.

If we limit our consideration to the on-axis pressure radiated by a circular piston, we may evaluate the integral analytically. The analysis follows similar steps to that for the frequency domain. Thus, a point on the surface is located by its polar coordinates $\left(R_{s}, \theta_{s}\right)$. The distance from such a point to the field point at distance $z$ from the surface is $r=\left(R_{s}^{2}+z^{2}\right)^{1 / 2}$. The value of $r$ is independent of $\theta_{s}$, so an integration over this variable leads to a $2 \pi$ factor. In addition, $v_{s}$ for a piston is the same for all $R_{s}$. Thus, Eq. (8.6.3) becomes

$$
\begin{equation*}
p\left(\bar{x}_{0}, t\right)=\rho_{0} \int_{0}^{a} \frac{1}{\left(R_{s}^{2}+z^{2}\right)^{1 / 2}} \frac{\partial}{\partial t} v_{s}\left(t-\frac{\left(R_{s}^{2}+z^{2}\right)^{1 / 2}}{c}\right) R_{s} d R_{s} \tag{8.6.5}
\end{equation*}
$$

At this stage, we change variables from $R_{s}$ to the radial distance $r$. Because $z$ is constant, differentiation of $r^{2}=R_{s}^{2}+z^{2}$ leads to $r d r=R_{s} d R_{s}$. Furthermore, the limits
$R_{s}=0$ and $R_{s}=a$ correspond to $r=z$ and $r=\left(a^{2}=z^{2}\right)^{1 / 2}$, which leads to

$$
\begin{equation*}
p\left(\bar{x}_{0}, t\right)=\rho_{0} \int_{z}^{\left(a^{2}+z^{2}\right)^{1 / 2}} \frac{\partial}{\partial t} v_{s}\left(t-\frac{r}{c}\right) d r \tag{8.6.6}
\end{equation*}
$$

The last step is to recognize the meaning of the derivative of $v_{s}$. Let $\Theta$ denote the argument, $\Theta=t-r / c$. Then, the chain rule for differentiation leads to

$$
\begin{gather*}
\frac{\partial}{\partial t} v_{s}\left(t-\frac{r}{c}\right) \equiv \frac{\partial \Theta}{\partial t} \frac{d}{d \Theta} v_{s}(\Theta)=\frac{d v_{s}}{d \Theta}  \tag{8.6.7}\\
\frac{\partial}{\partial r} v_{s}\left(t-\frac{r}{c}\right) \equiv \frac{\partial \Theta}{\partial r} \frac{d}{d \Theta} v_{s}(\Theta)=-\frac{1}{c} \frac{d v_{s}}{d \Theta}=-\frac{1}{c} \frac{\partial}{\partial t} v_{s}\left(t-\frac{r}{c}\right)
\end{gather*}
$$

Usage of this property to replace the time derivative with a derivative with respect to $r$ leads to

$$
\begin{equation*}
p(z, t)=-\rho_{0} c \int_{z}^{\left(a^{2}+z^{2}\right)^{1 / 2}} \frac{\partial}{\partial r} v_{s}\left(t-\frac{r}{c}\right) d r \tag{8.6.8}
\end{equation*}
$$

The integrand is a perfect differential, so we find that

$$
\begin{equation*}
p(z, t)=\rho_{0} c\left[v_{s}\left(t-\frac{z}{c}\right)-v_{s}\left(t-\frac{\left(a^{2}+z^{2}\right)^{1 / 2}}{c}\right)\right] \tag{8.6.9}
\end{equation*}
$$

This is the time-domain equivalent of Eq. (8.3.4). The first term is the center wave, which leaves the center of the piston at time $z / c$ prior to its arrival at the field point. The second term is the edge wave, whose propagation time prior to arrival at the field point is $\left(a^{2}+z^{2}\right)^{1 / 2} / c$.

Consider the case where the surface vibration is a pulse, such that $v_{s}(t)=0$ for $t<0$ or $t>T$. Then, the signal received at an axial point will be the same shape pulse in the edge wave, followed by the negative of the first pulse emanating from the edge of piston. The interval between the arrival times is $\left[\left(a^{2}+z^{2}\right)^{1 / 2}-z\right] / c$. If this interval is greater than $T$, the pulses will appear as separate events. This inequality may be solved for the range of $z$ in which the pulses occur individually. It is $z<\left(a^{2}-c^{2} T^{2}\right) /(2 c T)$.

Some insight into the farfield behavior may be gained by considering axial locations at which $z / a \gg 1$. Then, $\left(a^{2}+z^{2}\right)^{1 / 2} \approx z+a^{2} /(2 z)$. The second term is much less than $z$, so application of a Taylor series to Eq. (8.6.9) leads to

$$
\begin{equation*}
p(z, t) \rightarrow \frac{\rho_{0} a^{2}}{2 z} \dot{v}_{s}\left(t-\frac{z}{c}\right) \tag{8.6.10}
\end{equation*}
$$

Thus, the pressure signal received at an axial point far from the piston will mirror the surface acceleration.


Fig. 8.18 The pressure signal received at a point on the axis of a circular piston whose velocity $v_{s}$ is a square wave

An interesting consequence of this property is what is heard when a piston moves as a square wave, which was the situation in Example 8.5. No pressure is generated when the velocity is constant, so the only sound that reaches the farfield is a series of spikes at the discontinuities. This is illustrated in Fig. 8.18, which depicts the time-domain construction of $v_{s}$ and the pressure waveform at $z=4 \mathrm{~m}$, as opposed to the representation of these signals as truncated Fourier series in the aforementioned example. The value of $z / c=11.76 \mathrm{~ms}$, so the first impulse to be observed for $t>0$ is a negative one corresponding to the discontinuous decrease of $v_{s}$ at $t=-11.5 T$. Of course, it is not possible to change the velocity discontinuously, because doing so would require an infinite force. Thus, the graph is an approximation of a rapid transition, just as an impulsive force is an approximation of a large force that acts over a very short time interval.

Additional analyses have been performed for various off-axis locations. For example, Blackstock ${ }^{12}$ describes an approximate time-domain farfield analysis for a circular piston. Pierce ${ }^{13}$ shows how to determine the pressure radiated by a circular piston at an arbitrary field point as a single integral, rather than an integral over an area. These analyses are somewhat complicated, so we shall not explore them.

### 8.7 Homework Exercises

Exercise 8.1 A rectangular plate is flush mounted in an infinite rigid baffle that lies in a vertical plane $x y$. The plate is very rigid, so it undergoes a uniform translational vibration $v_{z}$ in the normal $z$ direction. The frequency of this vibration might be any value up to 1.2 kHz . Within this interval, the farfield mean-squared pressure in the (horizontal) $x z$ plane at $30^{\circ}$ from the $z$-axis should be no more than 6 dB less than the on-axis value. Also, the farfield mean-squared pressure at $5^{\circ}$ from the $z$-axis in the

[^29]vertical $y z$ plane should be no more than 6 dB less than the on-axis value. Determine the horizontal and vertical edge lengths that meet these specifications.


Exercise 8.1
Exercise 8.2 The sketch depicts a square plate that is embedded in a very large rigid baffle. The plate is flexible, with motion actuated $180^{\circ}$ out-of-phase from one corner to the next, with the result that the surface normal velocity is $v_{z}=V_{0}\left(x y / b^{2}\right) \cos (\omega t)$, where $b$ is the length of each edge. Derive an expression for the farfield directivity. For the case where $k b=10$, evaluate and graph this function for points in the $x z$ plane and $y z$ plane, as well as a plane that contains the $z$-axis at $45^{\circ}$ from the $x$ - and $y$-axes.


Exercise 8.2
Exercise 8.3 The sketch depicts a submerged square transducer mounted on a large rigid baffle that lies in the vertical plane. The edges of the transducer are horizontal and vertical, with the upper edge at depth $h$ below the free surface. This depth is sufficiently large that the baffle may be considered to extend infinitely in all directions for the purpose of formulating the Rayleigh integral. The transducer executes a uniform vibration $v_{0} \sin (\omega t)$ in the $z$ direction, which is perpendicular to the plane of the baffle. It is permissible to approximate the free surface as being pressurerelease for waves in the water. (a) Derive an expression for the complex amplitude
of the pressure in the farfield as a function of spherical coordinates for which $z$ is the polar axis and $r$ measures the distance from the $x y z$ origin. (b) Consider the case where $a=0.4 \mathrm{~m}$ and the frequency is 5 kHz . Determine and plot the farfield radiation $(r / a)\left|p_{\mathrm{ff}}\right| /\left(\rho 0 c v_{0}\right)$ as a function of angle from the $z$-axis for points in the $y z$ plane.


Exercise 8.3
Exercise 8.4 It is desired to build a baffled circular piston projector that generates a signal in air whose beamwidth is less than $1.5^{\circ}$ for frequencies greater than 16 kHz . What is the smallest radius $a$ of the piston that fits this specification? For this value of $a$, what is the beamwidth at $16 \mathrm{kHz}, 1.6 \mathrm{kHz}$, and 160 Hz ? Graph the farfield directivity at each of these frequencies.

Exercise 8.5 A Gaussian velocity distribution on a flat surface is one for which the velocity distribution depends on the polar distance $R_{s}$ from the center according to $v_{z}=V_{0} \exp \left(-\beta R_{s}^{2}\right) \sin (\omega t)$. This distribution is desirable because it limits diffraction effects that cause nulls due to a discontinuous surface velocity distribution. In reality, this concept is an approximation because $R_{s}$ cannot exceed the active radius $a$. Thus, the Gaussian surface velocity distribution applies for $R_{s}<a$, and $v_{z}=0$ for $R_{s}>a$. Use numerical methods to evaluate the farfield directivity when $k a=12$. Perform this evaluation for $\beta=0, \beta=0.5, \beta=2$, and $\beta=8$. What conclusions can be drawn from these evaluations?

Exercise 8.6 Derive an expression for the farfield directivity of the transducer in Example 8.3. Evaluate the result for $k a=5$ and $k a=20$. Compare each pattern to the directivity of a baffled piston.

Exercise 8.7 A circular plate flush mounted in an infinite baffle executes a harmonic motion in which it rotates as a rigid body about a diametral line. The consequence is that the surface velocity is $v_{1}(y / a) \sin (\omega t)$. Derive an expression for the farfield pressure. Evaluate and graph this function for points in the $x z$ plane and $y z$ plane for the case where $k a=10$.


Exercise 8.7
Exercise 8.8 Drawing I shows a projector embedded in the surface of a rigid infinite baffle. The projector face is a semicircular piston, with the velocity being constant along the face. If $R_{s}, \theta_{s}$ are polar coordinates for a point on the face, with $x=$ $R_{s} \cos \left(\theta_{s}\right), y=R_{s} \sin \left(\theta_{s}\right)$, then a harmonic oscillation of the piston leads to a surface velocity described by

$$
v_{z}=\left\{\begin{array}{l}
v_{0} \cos (\omega t), \quad 0<R_{s}<a, \quad-\pi / 2<\theta_{s}<\pi / 2 \\
0 \text { otherwise }
\end{array}\right.
$$

(a) Consider a field point $\bar{x}_{0}$ on the $z$-axis $(\psi=0)$. Starting with the general Rayleigh integral, derive an expression for the complex amplitude of the pressure at such a location. How does that result compare to the result for the case of a full circular piston? (b) Starting with the general Rayleigh integral, derive an expression for the complex amplitude of the pressure at location $\bar{x}_{0}$ in the farfield, $r \gg a$. The final result may be left in integral form. (c) In Drawing II, an infinite baffle is situated in the vertical plane. A vibrating piston is mounted on the baffle such that its center is at the free surface of the water. For the purpose of the analysis, the free surface of the water may be taken to be pressure-release for waves in the water and rigid for waves in the air. Let $P_{b}(r, \phi, \theta)$ denote the pressure derived in Part (b). Describe in terms of $P_{b}(r, \phi, \theta)$ the pressure in the water and in the air.


Exercise 8.8

Exercise 8.9 A circular piston in a rigid baffle executes a transverse vibration $v_{0} \sin (\omega t)$ everywhere along its surface. The following facts are known: (1) The fluid is water. (2) The frequency is 50 kHz . (3) The first null in the farfield pressure occurs at $3^{\circ}$ from the $z$-axis. (4) The pressure at 3 m along the $z$-axis from the plate is 14 kPa . Determine the radius $a$ and velocity amplitude $v_{0}$ that correspond to the stated conditions.

Exercise 8.10 An ultrasonic piston transducer in a large rigid baffle transmits a signal into water. A scan of the axial field indicates that the largest pressure is 195 $\mathrm{dB} / / 1 \mu \mathrm{~Pa}$, and that farthest distance at which it occurs is 28 mm from the face of the piston. It also is observed that the farthest antinode occurs at 12 mm from the face. Determine the radius $a$ of the piston, the frequency, and the amplitude of the piston's velocity. Also estimate the minimum radial distance at which farfield properties can be expected to be observed.
Exercise 8.11 A transducer is composed of multiple piezoeceramic elements that are driven in-phase at frequency $\omega$, with the individual elements approximately filling a circle of radius $a$. This circle is surrounded by a rigid baffle, that is, sufficiently large to consider it to be infinite. Because this transducer is not exactly a circular piston the effective radius $a_{\text {eff }}$ for acoustical properties might not be exactly the geometrical radius $a$. The value of $a_{\text {eff }}$, as well as the local speed of sound, can be determined from the axial distances of the farthest two minima of the on-axis pressure amplitude. Let these distances be $z_{A}$ and $z_{B}<z_{A}$. (a) Derive formulas for $a_{\text {eff }}$ and $c$ corresponding to a known frequency $\omega$ and measured values of $z_{A}$ and $z_{B}$. (b) Evaluate the results in Part (a) for the case of a 10 kHz signal with $z_{A}=0.395 \mathrm{~m}$, and $z_{B}=0.105 \mathrm{~m}$. (c) How many pressure antinodes will be observed on the symmetry axis when the parameters are those in Part (b)? (d) For the parameters in Part (b), what distance is a good estimate for the onset of farfield behavior? (e) What is the beamwidth corresponding to the parameters in Part (b)?

Exercise 8.12 An underwater baffled piston projector must have a maximum sound pressure level of $240 \mathrm{~dB} / / 1 \mu \mathrm{~Pa}$ on-axis when the frequency is 8 kHz . The farthest distance at which this maximum occurs should be 3 meters. (a) What is the required radius of the piston? (b) At what distance is it reasonable to consider farfield spreading to begin? (c) How many nulls are there on-axis? (d) What is the sound pressure level at 30 m from the projector on-axis? (e) What is the beamwidth of this projector? (f) How many nulls occur in the directivity?

Exercise 8.13 An infinite plane is the boundary of a half space filled with an ideal fluid. The entire plane outside a circle of radius $a$ executes an in-phase harmonic vibration normal to the plane. With $z$ being the axial direction for a set of cylindrical coordinates, and $z=0$ designated as the plane, the surface velocity is $v_{z}=v_{0} \sin (\omega t)$ if $R>a$ and $v_{z}=0$ if $R<a$. Derive expressions for the pressure along the z -axis and the pressure at an arbitrary location in the farfield. For the case where $k a=20$, graph the axial distribution $|P| /\left(\rho_{0} c v_{0}\right)$ as a function of $z / a$. Also, graph the farfield pressure $\left|P_{f f}\right| /\left(\rho_{0} c v_{0}\right)$ as a function of distance $R / a$ from the symmetry axis. The axial distance for this evaluation should be $z=R_{0}$.

Exercise 8.14 A projector in an infinite baffle consists of a ring whose inner radius is $a / 2$ and whose outer radius is $a$. The complex normal velocity amplitude at frequency $\omega$ is $v_{0}$ over the surface of the ring, while the region inside the inner radius is motionless. Derive expressions for the farfield pressure and the pressure on-axis. Given that $k a=25$, evaluate $(r / a)\left|p_{\mathrm{ff}} /\left(\rho_{0} c v_{0}\right)\right|$ as a function of the polar angle, and the axial pressure distribution $\left|p(\psi=0) /\left(\rho_{0} c v_{0}\right)\right|$ as a function of distance from the piston. Do nulls occur in either pattern?

Exercise 8.15 Example 8.2 analyzed the farfield radiation of a membrane transducer. The membrane's surface velocity is well approximated be a parabolic distribution, $V_{s}=v_{1}\left(1-R^{2} / a^{2}\right) \sin (\omega t)$ if $R<a, V_{s}=0$ if $R \geq a$. Evaluate $\left|p /\left(\rho_{0} c v_{0}\right)\right|$ as a function of distance along the axial center line from the plane of the baffle. Perform this evaluation for $k a=3$ and 20.

Exercise 8.16 The sketch depicts a transducer that radiates sound into water. It consists of a piston that is flush mounted in an infinite baffle. The piston, whose mass is $M=0.6 \mathrm{~kg}$, is restrained by spring $K$. The displacement at the other end of the spring is controlled by an actuator that can exert whatever force is required to impose a specified harmonic motion $v_{\text {act }} \cos (\omega t)$. Design specifications are as follows: (1) The piston radius is 150 mm . (2) The operational frequency is 2 kHz . (3) The natural frequency of the mass-spring system in a vacuum equals the operational frequency. (4) The maximum pressure on-axis at the operational frequency is 200 kPa . Given these requirements, determine $v_{\text {act }}$ and the force exerted by the fluid on the piston.


Exercise 8.16 and 8.17.
Exercise 8.17 The sketch depicts a piston transducer that is flush mounted in an infinite baffle. The piston, whose mass is $M$, is restrained by spring $K$. The displacement at the other end of the spring is controlled by an actuator that can exert whatever force is required to impose a specified harmonic motion $\operatorname{Re}\left(V_{\text {act }} \exp (i \omega t)\right)$. The task is to determine radiation properties as a function of the frequency. Such an analysis is initiated by deriving a differential equation of motion for $v_{\text {piston }}$ that accounts for the fluid loading acting on the piston. Solution of this equation for the specified actuator motion will yield $\left|V_{\text {piston }} / V_{\text {act }}\right|$ as a function of $\omega$. From that result evaluate the timeaveraged radiated power $\mathcal{P}_{\text {av }} /\left(\rho_{0} c\left|V_{\text {act }}\right|^{2}\right)$. Graph both quantities as functions of $\omega$ in the range from zero to $20 \%$ greater than the natural frequency of the mass-spring
system in a vacuum. Carry out the analysis for the case where the fluid is water, then repeat the computation for the case where the fluid is air. The transducer's properties are $a=150 \mathrm{~mm}, M=0.2 \mathrm{~kg}$, and $K=45\left(10^{6}\right) \mathrm{N} / \mathrm{m}$.

Exercise 8.18 A piston, whose radius is 250 mm , is embedded in an infinite rigid baffle. The fluid is air. The piston was at rest until $t=0$, at which instant it begins a parabolic pulse, in which the surface velocity is $v_{s}=0.3\left(t / T-t^{2} / T^{2}\right) \mathrm{m} / \mathrm{s}$ if $0<t<T$. Evaluate the waveform at the on-axis positions $z=0,4 a$, and $32 a$ in the case where $T=0.25 a / c$. Compare each to the limiting form in Eq. (8.6.10). For each location, determine the time required for the signal to attain a steady state.

Exercise 8.19 The waveform plotted below is the normal velocity of a circular piston in an infinite baffle. It consists of a single cycle of a sawtooth. Evaluate and plot $p /\left(\rho_{0} c v_{0}\right)$ as a function of the nondimensional retarded time $(c / a)(t-z / c)$ for on-axis locations. Perform the calculation for $z=a$ and $z=4 a$, with the pulse interval set at $T=0.3 a / c$ and $T=3 a / c$.


Exercise 8.19

## Chapter 9 <br> Modal Analysis of Waveguides

We return to situations where boundaries guide the direction in which sound may propagate. The difference from the one-dimensional waveguides considered in Chaps. 2 and 3 is that some feature will vary transversely to the propagation direction. It might be that the excitation varies in the transverse direction, which is analogous to the effect of a spherical source that is not radially symmetric. This situation commonly is encountered in turbine engines and ocean acoustics. Another cause for transverse variation of a propagating signal is a nonconstant spacing between the confining boundaries. Examples of this situation are musical horns and some speaker systems. Both features often are encountered in HVAC systems. Fully realistic models that account for all effects are difficult to analyze, so they are typically addressed with computer methods. However, if we make some idealizations, there are a number of configurations that are amenable to analysis. The studies that follow provide much understanding of basic phenomena, and all are descriptive of situations that one might encounter.

### 9.1 Propagation in a Horn

Unlike the one-dimensional waveguides considered previously, a horn has a cross section that is not constant along the axis of propagation. For example, conical horns, such as those commonly used by the coxswain in shell rowboat racing, have a cross-sectional radius that is proportional to the axial distance $x$. A configuration of particular interest is an exponential horn, in which the cross-sectional area depends exponentially on $x$. In principle, the shape of the cross section could change, but such configurations are rare. In any event, the actual shape of the cross section will be irrelevant to the analysis.

The idea of attaching a horn to a small source is to attain some features of an unadorned large source. One such feature is directivity. We know that a source whose
size is small relative to a wavelength will radiate almost omnidirectionally. The large area of a horn at its open end has the effect of increasing the size of the radiator. This aspect may be recognized by considering a conical horn. In Fig.9.1a, a spherical cap in a baffle executes a purely radial vibration. A rigid cone is mounted on the perimeter of the cap, with the apex of the cone coincident with the center of the sphere. In Fig. 9.1b, the cone is positioned on a freely standing sphere of the same radius that executes the same radially symmetric vibration everywhere on its surface. The particle velocity in the signal radiated by the sphere in case (b) is purely radial. This is compatible with the rigidity of the cone, which requires that the particle velocity normal to its surface be zero. It follows that the field inside the cone is the same as what it would be if the cone were not present. (This assertion assumes that there is no reflection from the cone's open end.) Furthermore, what occurs outside the cone is irrelevant to the field inside the cone. This leads to recognition that the field in case (a) is a radially symmetric spherical wave within the cone. Thus, the presence of the conical horn in case (a) changes the effective size of the radiator from the diameter $2 a_{0}$ of the cap to the diameter $2 a_{1}$ of the opening. It also makes it unnecessary to mount the cap in the baffle.

(b)


Fig. 9.1 A conical horn mounted on a spherical transducer. In case a, the transducer is a cap in a baffle, whereas in case $\mathbf{b}$, it is in free space

Another reason to mount a horn on a source pertains to applications in which it is desired to have a constant output over a broad frequency range. Without the horn, attainment of this objective would require either a small transducer whose velocity amplitude increases inversely to the frequency, or else a very large transducer. Neither feature is desirable for reasons that will be disclosed when we see why a horn avoids these issues.

### 9.1.1 The Webster Horn Equation

The model to be constructed is based on a fundamental assumption: The pressure perturbation depends only on the axial position $x$ and $t$. A corollary that follows from Euler's equation is that the particle velocity is solely in the axial direction, with $v_{x}$ also being a function of $x$ and $t$. Clearly, the latter cannot be exactly correct, because if it were there would be no signal in the region into which the horn expands.

Intuitively, it is reasonable to expect that this assumption is valid if the cross-sectional area is a function $\mathcal{A}(x)$ that varies slowly relative to the scale of a wavelength. The correctness of this assumption is an item we will explore.

The field equation is obtained by enforcing conservation of mass and the momentum-impulse principle. These considerations parallel those that led to the one-dimensional wave equation in Chap. 2. However, rather than using finite-sized control volumes and finite increments in time, here we shall take a less rigorous, but more expedient, approach that considers the rates of change in the state of differential control volumes. We begin with conservation of mass for the stationary control volume in Fig. 9.2.

Fig. 9.2 A differential stationary control volume used to describe conservation of mass in a horn


The wall of the horn forms the sides of this region. It is impermeable, so there is no mass flow across the sides of the control volume. Thus, the net rate at which mass flows into the control volume over the two cross sections must equal the rate at which the mass of this control volume increases. For the cross section at $x$, positive $v_{x}$ transports mass into the control volume, and $\rho v_{x}$ evaluated at $x$ is the mass flow rate per unit area. Positive $v_{x}$ on the face at $x+d x$ transports mass out of the control volume, and $\rho v_{x}$ evaluated at $x+d x$ is the rate per unit area there. According to the central limit theorem, $\rho(x, t) \mathcal{A}(x) d x$ is the mass of the control volume at time $t$. Thus, it must be that
$\rho(x, t) v_{x}(x, t) \mathcal{A}(x)-\rho(x+d x, t) v_{x}(x+d x, t) \mathcal{A}(x+d x)=\frac{\partial}{\partial t}[\rho(x, t) \mathcal{A}(x) d x]$

Division by $d x$ converts the left side of the preceding to a partial derivative of $\rho v_{x} \mathcal{A}$, so the expression becomes

$$
\begin{equation*}
\mathcal{A} \frac{\partial \rho}{\partial t}+\frac{\partial}{\partial x}\left(\rho v_{x} \mathcal{A}\right)=0 \tag{9.1.2}
\end{equation*}
$$

The second term is nonlinear because it is the product of two state variables. A linearized analysis restricts the particle velocity to being a small quantity (relative to the speed of sound), and the density is restricted to being little changed from the ambient value $\rho_{0}$. Correspondingly, we linearize the continuity equation by replacing
$\rho v_{x}$ with $\rho_{0} v_{x}$. For a homogeneous medium, $\rho_{0}$ is the same at all locations, so the linearized continuity equation for a horn is

$$
\begin{equation*}
\mathcal{A} \frac{\partial \rho}{\partial t}+\rho_{0} \frac{\partial}{\partial x}\left(\mathcal{A} v_{x}\right)=0 \tag{9.1.3}
\end{equation*}
$$

If the cross section is constant, this relation reduces to the continuity equation for truly one-dimensional plane waves.

The momentum equation is derived by considering a control volume that moves with the particles. Because it tracks a specific set of particles, its selection simplifies application of the momentum-impulse principle for a system of particles. The control volume in Fig. 9.3a consists of the fluid contained between cross sections at $x$ and $x+d x$ at time $t$. The forces depicted there consist of the pressure resultant acting on each cross section and the distributed force exerted by the wall of the horn in reaction to the pressure the fluid exerts in the wall. It would be awkward to account for the axial component of the wall force.


Fig. 9.3 Control volume used to form the derivative form of the momentum-impulse principle. The domain moves in unison with the particles contained between two adjacent cross sections. a The domain consists of all particles contained between the adjacent cross sections. $\mathbf{b}$ The domain is a cylinder whose cross section is constant

The control volume in Fig. 9.3b is more amenable to the analysis. It is a general cylinder whose cross section is $\mathcal{A}(x)$. This shape exploits the fact that shear stresses are negligible for an ideal fluid. Thus, the force system exerted on the sides of this cylindrical domain is the stress resultant of the pressure exerted by the adjacent fluid. This resultant acts perpendicularly to the cylindrical surface, so it has no component in the axial direction. (The transverse resultant actually must be zero for the assumed motion, because there is no acceleration in that direction.)

There is no evidence of a variable cross section in the second force diagram, as in Fig. 9.3b. Thus, the time-domain Euler equation, which is obtained from conservation of momentum, may be applied to the motion in the $x$ direction

$$
\begin{equation*}
\rho \frac{d}{d x} v(x, t)=-\frac{\partial p}{\partial x} \tag{9.1.4}
\end{equation*}
$$

A linearized equation of motion results from dropping any terms that are products of small variables. Thus, the acceleration is approximated as a partial time derivative, and the current density is approximated by its ambient value. Doing so reduces the equation of motion for the control volume to

$$
\begin{equation*}
\rho_{0} \frac{\partial v_{x}}{\partial t}=-\frac{\partial p}{\partial x} \tag{9.1.5}
\end{equation*}
$$

This is the same as the $x$ component of Euler's equation.
Equations (9.1.3) and (9.1.5) constitute two equations for three unknowns: $p, \rho$, and $v_{x}$. The third relation is the equation of state, whose linearized form is $p=$ $c^{2}\left(\rho-\rho_{0}\right)$. We use this equation to eliminate $\rho$ from the continuity equation. To eliminate $v_{x}$, we multiply Eq. (9.1.5) by $\mathcal{A}$ and then differentiate it with respect to $x$. Because $\mathcal{A}$ is independent of $t$, the result may be written as

$$
\begin{equation*}
\rho_{0} \frac{\partial^{2}}{\partial x \partial t}\left(\mathcal{A} v_{x}\right)=-\frac{\partial}{\partial x}\left(\mathcal{A} \frac{\partial p}{\partial x}\right) \tag{9.1.6}
\end{equation*}
$$

Differentiation of Eq. (9.1.3) with respect to $t$ yields the same $v_{x}$ term,

$$
\begin{equation*}
\frac{\mathcal{A}}{c^{2}} \frac{\partial^{2} p}{\partial t^{2}}+\rho_{0} \frac{\partial^{2}}{\partial x \partial t}\left(\mathcal{A} v_{x}\right)=0 \tag{9.1.7}
\end{equation*}
$$

Elimination of the mixed derivative term in these two equations yields the Webster horn equation

$$
\begin{equation*}
\frac{1}{\mathcal{A}} \frac{\partial}{\partial x}\left(\mathcal{A} \frac{\partial p}{\partial x}\right)-\frac{1}{c^{2}} \frac{\partial^{2} p}{\partial t^{2}}=0 \tag{9.1.8}
\end{equation*}
$$

This partial differential equation has variable coefficients except in the case of a uniform bar. In the frequency domain, it becomes an ordinary differential equation,

$$
\begin{equation*}
p=\operatorname{Re}\left(P(x) e^{i \omega t}\right) \Longrightarrow \frac{d^{2} P}{d x^{2}}+\frac{1}{\mathcal{A}}\left(\frac{d \mathcal{A}}{d x}\right)\left(\frac{d P}{d x}\right)+k^{2} P=0 \tag{9.1.9}
\end{equation*}
$$

This too is referred to as the Webster horn equation. There are horn shapes $\mathcal{A}(x)$ for which it is possible to obtain an analytical solution in the frequency domain. We also will develop an approximate solution for arbitrary $\mathcal{A}(x)$. However, before we pursue these tasks, let us return to the conical horn because doing so will allow us to identify when the Webster horn equation may be used.

Figure 9.4 depicts a conical waveguide whose semi-vertex angle is $\beta$. An arbitrary point is located by its cylindrical coordinates $(x, R)$ and its spherical coordinates

Fig. 9.4 Cylindrical and spherical coordinates of a field point within a conical horn

$(r, \psi)$, with both coordinate systems defined relative to the apex of the cone. The spherical cap at the left executes a uniform radial vibration at frequency $\omega$.

From the discussion of Fig. 9.1, we know that the field is a radially symmetric wave, so an outgoing wave is

$$
\begin{equation*}
p_{\text {true }}=\operatorname{Re}\left(\frac{D}{r} e^{-i k r} e^{i \omega \tau}\right) \tag{9.1.10}
\end{equation*}
$$

An analytical solution of the Webster horn equation for this case is available. Let us compare it to the spherical wave solution in order to grasp the effect of the approximations embedded in that equation. The cross-sectional radius of the waveguide is $x \tan \beta$, which leads to $A(x)=\pi(\tan \beta)^{2} x^{2}$. The Webster horn equation in this case is

$$
\begin{equation*}
\frac{\partial^{2} p}{\partial x^{2}}+\frac{2}{x} \frac{\partial p}{\partial x}-\frac{\partial^{2} p}{\partial t^{2}}=0 \tag{9.1.11}
\end{equation*}
$$

If, instead of $x$, the independent variable were $r$, this would be the wave equation for a radially symmetric spherical wave. Thus, if we ignore reflections from the far end of the horn, the solution of the Webster horn equation for an outgoing wave is

$$
\begin{equation*}
p_{\text {horn }}=\operatorname{Re}\left(\frac{D^{\prime}}{x} e^{-i k x} e^{i \omega t}\right) \tag{9.1.12}
\end{equation*}
$$

The radial distance to a point on a specific cross section is $r=x / \cos \psi$. Setting $D^{\prime}=D$ makes this solution and the spherical coordinate solution, Eq. (9.1.10), the same on the axis, $\psi=0$.

With increasing distance from the axis, the solutions diverge. The relative error in the amplitude at a specific $\psi$ is obtained by replacing $r$ in Eq. (9.1.10) with $x / \cos \psi$, which leads to

$$
\begin{equation*}
\varepsilon_{\text {amplitude }} \equiv \frac{\left|P_{\text {horn }}\right|-\left|P_{\text {true }}\right|}{\left|P_{\text {true }}\right|}=\cos \psi-1 \tag{9.1.13}
\end{equation*}
$$

Another aspect is the phase error, which must be small if the waveform is to be positioned correctly. This quantity is

$$
\begin{equation*}
\varepsilon_{\text {phase }}=k x-k r=k x\left(1-\frac{1}{\cos \psi}\right) \tag{9.1.14}
\end{equation*}
$$

Both errors are underestimates that increase monotonically with increasing $\psi$. The largest errors at fixed $x$ occur at the wall, where $\psi=\beta$. A trigonometric identity is $1-\cos \beta \equiv(\sin \beta) \tan (\beta / 2)$. The resulting error estimates are

$$
\begin{equation*}
\varepsilon_{\text {amplitude }}=-\sin \beta \tan \frac{\beta}{2}, \quad \varepsilon_{\text {phase }}=-k x \tan \beta \tan \frac{\beta}{2} \tag{9.1.15}
\end{equation*}
$$

The amplitude error will be small if $\beta \ll 1$, that is, if the cross section expands slowly. However, even if the amplitude error is acceptable, the fact that $\varepsilon_{\text {phase }}=$ $(k x / \cos \beta) \varepsilon_{\text {amplitude }}$ means that the amplitude and phase error will have comparable magnitude only if $k x$ is not large compared to unity. In other words, $\beta$ should be much less than 1 rad and $x$ should not be much larger than a wavelength. Let us assess these criteria for the audible range of a horn in air. Consider $\beta=10^{\circ}$. The amplitude error at $\psi=\beta$ is $1.5 \%$. At the low end of the audible spectrum, 20 Hz , the phase error will be less than $10 \%$ if $x<17.5 \mathrm{~m}$. In contrast, at 10 kHz , the phase error will be less than $10 \%$ if $x<35 \mathrm{~mm}$. The low-frequency case suggests that an English horn or euphonium might be well described by the Webster horn equation, but this model might not be adequate for a flute or piccolo.

These criteria may be generalized to any shape horn. Let $R_{\text {wall }}(x)$ be the transverse distance to the wall at a specified axial position. For a cone, we have $d R_{\text {wall }} / d x=\tan \beta$, so small $\beta$ corresponds to small values of $d R_{\text {wall }} / d x$ and $\tan (\beta / 2) \approx(1 / 2) d R_{\text {wall }} / d x$. Thus, the amplitude and phase errors will not be excessive if $d R_{\text {wall }} / d x \ll 1$ and $k x$ is not much larger than a wavelength.

### 9.1.2 Exponential Horn

Let us examine whether it is possible for a simple plane wave to exist within a horn. Let $\kappa$ be the constant wavenumber. Then, substitution of

$$
\begin{equation*}
P=B e^{-i \kappa x} \tag{9.1.16}
\end{equation*}
$$

into the frequency-domain Webster horn equation requires that

$$
\begin{equation*}
\left(k^{2}-\kappa^{2}\right) \mathcal{A}-i \kappa \frac{d \mathcal{A}}{d x}=0 \tag{9.1.17}
\end{equation*}
$$

The wavenumber of a simple plane wave is a real quantity. Because $\mathcal{A}$ is real, it is evident that in no circumstance, other than constant $\mathcal{A}$, is it possible for a simple wave to exist in a horn.

In the special case where $d \mathcal{A} / d x$ is proportional to $\mathcal{A}$, that is, $d \mathcal{A} / d x=b \mathcal{A}$, the area $\mathcal{A}$ factors out of Eq. (9.1.17). What remains is a complex quadratic equation for $\kappa$. The cross-sectional variation fitting this description constitutes an exponential horn

$$
\begin{equation*}
\mathcal{A}(x)=A_{0} e^{b x} \tag{9.1.18}
\end{equation*}
$$

The growth factor $b$ may be negative, corresponding to a cross section whose radius decreases with increasing $x$. In most cases, the cross section is either rectangular or circular, with the shape held constant at all $x$, but these details are irrelevant to the analysis.

When the cross-sectional area varies exponentially, Eq. (9.1.17) leads to a quadratic equation for the wavenumber, specifically

$$
\begin{equation*}
\kappa^{2}+i b \kappa-k^{2}=0 \tag{9.1.19}
\end{equation*}
$$

If $k>|b| / 2$, there are two roots complex roots. The one whose real part is positive corresponds to propagation in the positive $x$ direction. It is

$$
\begin{equation*}
\kappa=\left(k^{2}-\frac{b^{2}}{4}\right)^{1 / 2}-i \frac{b}{2} \text { if } k>\frac{|b|}{2} \tag{9.1.20}
\end{equation*}
$$

The other root corresponds to waves that propagate in the direction of decreasing $x$. It is

$$
\begin{equation*}
\kappa^{\prime}=-\left(k^{2}-\frac{b^{2}}{4}\right)^{1 / 2}-i \frac{b}{2} \text { if } k>\frac{|b|}{2} \tag{9.1.21}
\end{equation*}
$$

Both waves will exist if there are reflections at one end. The imaginary part of both $\kappa$ and $\kappa^{\prime}$ correspond to waves that decrease with increasing $x$ if $b$ is positive. As was noted above, a horn whose cross section decreases with increasing $x$ may be represented by a negative $b$. The alternative is to reverse the $x$-axis, so that it is oriented in the direction of increasing cross section. In either view, the amplitude decreases for a wave that propagates in the sense of increasing area, and it increases for a wave that propagates in the sense of decreasing area. The general solution when $k>|b| / 2$ is

$$
\begin{equation*}
P=e^{-b x / 2}\left(B_{1} e^{-i\left(k^{2}-b^{2} / 4\right) x}+B_{2} e^{i\left(k^{2}-b^{2} / 4\right) x}\right) \text { if } k>|b| / 2 \tag{9.1.22}
\end{equation*}
$$

It is possible that $k$ is smaller than $b / 2$, in which case the square root yields an imaginary value. Both wave numbers are negative imaginary values

$$
\begin{equation*}
\kappa=-i\left[ \pm\left(\frac{b^{2}}{4}-k^{2}\right)^{1 / 2}+\frac{b}{2}\right] \text { if } k<\frac{|b|}{2} \tag{9.1.23}
\end{equation*}
$$

In this situation the waves that travel back and forth are evanescent waves. Neither solution corresponds to a wave that propagates. Rather, both are standing waves in which $|P|$ decreases on the sense of increasing values of $b x$. The evanescent field may be written as

$$
\begin{equation*}
P=e^{-b x / 2}\left(B_{1} e^{+\left(b^{2} / 4-k^{2}\right) x}+B_{2} e^{-\left(b^{2} / 4-k^{2} x\right)}\right) \tag{9.1.24}
\end{equation*}
$$

The case where $k=|b| / 2$ marks a transition from a propagating wave to evanescence. This condition marks the cutoff frequency,

$$
\begin{equation*}
\omega_{\text {cutoff }}=b c / 2 \tag{9.1.25}
\end{equation*}
$$

Below the cutoff frequency, a waveguide is seldom of interest because the timeaveraged power transported through the horn is essentially zero. (Power flow will be examined shortly.) The parameter $\kappa$ is the complex wavenumber. Its real part gives the wavelength. It also defines the phase speed, according to

$$
\begin{equation*}
c_{\text {phase }} \equiv \frac{\omega}{|\operatorname{Re}(\kappa)|}=\frac{c}{\left(1-b^{2} / 4 k^{2}\right)^{1 / 2}} \tag{9.1.26}
\end{equation*}
$$

Because the phase speed is a function of frequency, a signal consisting of more than one frequency will spread out, or disperse, as it propagates. For this reason, a relation like Eq. (9.1.26) is said to be a dispersion equation. However, the nature of other systems is such that it might not be possible to obtain a closed form solution for $\operatorname{Re}(\kappa)$ as a function of frequency. Thus, any equation, such as Eq. (9.1.20), that relates the complex wavenumber to the frequency will be said to be a dispersion equation.

In terms of dimensionless quantities, $c_{\text {phase }} / c$ for an exponential horn depends only on $k / b$, which is the relation depicted in Fig. 9.5. At high frequencies, the phase speed is close to the plane wave speed. Decreasing the frequency causes $c_{\text {phase }}$ to increase, until it becomes infinite at the cutoff frequency. Below $k=b / 2$, the phase speed is undefined because $\kappa$ is imaginary.


Fig. 9.5 Dispersion curve for an exponential horn

The cutoff effect results from the fact that $|\kappa|=k$ for any $|b|<2 k$. Consider a variety of waveguides at a specific frequency. As $b$ increases from a small value with $k$ fixed, it is necessary that $\operatorname{Im}(\kappa)$ increase in order to match $b$. This requires that $\operatorname{Re}(\kappa)$ decrease. At the cutoff condition, $b=2 k$, the imaginary part equals $k$, so $\operatorname{Re}(\kappa)=0$.

The simpler alternative when the larger end is excited is to let $b$ be negative, rather than reversing the sense of $x$. Doing so allows us to set $x=0$ as the driven end in any case. In the absence of reflections at the undriven end, setting $b<0$ leads to $|P(x=L) / P(x=0)|=e^{-b L}>1$. The signal at the exit would be substantially enhanced if $-b L \gg 1$. Such a configuration was used for antique hearing aids, as well as in the apparatus that recorded sound on a phonograph platter. A horn amplified the signal that actuated the encoding stylus.

The usual situation is that the particle velocity at the end $x=0$ is specified. Let us assume that the far end is terminated in a manner that prevents reflection. Accordingly, we set $B_{2}=0$ in Eq. (9.1.22). A general expression for the velocity may be obtained by substituting the pressure ansatz, Eq. (9.1.16), into Euler's equation. Doing so gives

$$
\begin{equation*}
V_{x}=\frac{\kappa}{\omega \rho_{0}} B e^{-i \kappa x} \equiv \frac{\kappa}{k} \frac{P}{\rho_{0} c} \tag{9.1.27}
\end{equation*}
$$

The expressions for $P(x)$ and $V(x)$ are general, but the behavior is fundamentally different above and below the cutoff frequency. Above cutoff, $\kappa$ is complex. This means that there is a frequency-dependent phase difference between the velocity and the pressure. Below the cutoff frequency, $\kappa$ is imaginary. This corresponds to frequency-independent $90^{\circ}$ phase difference between the particle velocity and pressure.

Satisfying the boundary condition will set the coefficient $B$. We equate $V_{x}$ at $x=0$ from the preceding to $V_{0}$. The value of $B$ is set by matching the particle velocity according to this relation to the complex amplitude $V_{0}$ of the velocity at that end. The result is

$$
\begin{equation*}
P(x)=\rho_{0} c V_{0} \frac{k}{\kappa} e^{-i \kappa x} \tag{9.1.28}
\end{equation*}
$$

The development began with two reasons horns are used presently. If the horn is sufficiently long, the open aperture will be much larger than the size of the opening where the horn is excited, thereby enhancing directivity of the signal that radiates from that end. However, the signal at the aperture is much weaker due to spreading loss, so the advantages of a horn regarding radiated power are not evident. An analysis of power flow begins with the observation that there is no variation over a cross section. Consequently, the power flowing across any cross section is the product of the intensity and the area

$$
\begin{align*}
\mathcal{P} & =\frac{1}{4} \mathcal{A}\left(P e^{i \omega t}+P^{*} e^{-i \omega t}\right)\left(V_{x} e^{i \omega t}+\text { c.c. }\right) \\
& =\frac{1}{2} \mathcal{A} \operatorname{Re}\left(P V_{x} e^{2 i \omega t}+P^{*} V_{x}\right) \tag{9.1.29}
\end{align*}
$$

This expression is applicable to any system that fits the assumptions embedded in the Webster horn equation. In the special case of an exponential horn, substitution of $P(x)$ and $\mathcal{A}(x)$ into this expression gives

$$
\begin{equation*}
\mathcal{P}=\frac{1}{2} \mathcal{A}_{0} \rho_{0} c \operatorname{Re}\left[\left(V_{0}\right)^{2} \frac{k}{\kappa} e^{2 i \omega t} e^{(-2 i \kappa+b) x}+\left|V_{0}\right|^{2} \frac{k}{\kappa} e^{-i\left(\kappa+\kappa^{*}-b\right) x}\right] \tag{9.1.30}
\end{equation*}
$$

The complex wavenumber is given in Eq. (9.1.20). If $\omega<\omega_{\text {cut }}$, then $\kappa$ is negative imaginary, which leads to

$$
\begin{equation*}
\mathcal{P}=\frac{1}{2} \mathcal{A}_{0} \rho_{0} c \operatorname{Re}\left[\left(V_{0}\right)^{2} \frac{k}{\kappa} e^{2 i \omega t}\right] e^{-\left(b^{2}-4 k^{2}\right)^{1 / 2} x} \text { if } k<b / 2 \tag{9.1.31}
\end{equation*}
$$

Thus, for frequencies below cutoff, the time-averaged power flow is zero, and the oscillatory part of the power flow decays exponentially with increasing distance. Above the cutoff frequency, $\kappa$ is complex. A few manipulations show that the power flow in this case is

$$
\begin{equation*}
\mathcal{P}=\frac{1}{2} \mathcal{A}_{0} \rho_{0} c \operatorname{Re}\left[\left(V_{0}\right)^{2} \frac{k}{\kappa} e^{2 i \omega\left(t-x / c_{\text {phase }}\right)}+\left|V_{0}\right|^{2} \frac{k}{\kappa}\right] \text { if } k>b / 2 \tag{9.1.32}
\end{equation*}
$$

This relation indicates that above the cutoff frequency, the time-averaged power flow is independent of the cross section's location. This property is predicted by the general theorem regarding power, which tells us that the average power that flows through the waveguide must be the same at all cross sections because energy does not flow out of the walls. The fluctuating part of the power flow propagates downstream as a second harmonic at the phase speed of the pressure wave.

The invariance of the time-averaged power means that we only need know the amount that is input at $x=0$. This is true regardless of the shape of the horn. For this reason, a quantity of interest for any type of horn is the ratio of $P$ at $x=0$ to the particle velocity $V_{0}$ at that end. This ratio is the throat impedance,

$$
\begin{equation*}
Z_{\text {throat }} \equiv \frac{\left.P\right|_{x=0}}{V_{0}} \tag{9.1.33}
\end{equation*}
$$

If we know this quantity, we may evaluate the time-averaged power input to the horn by the transducer according Eq. (9.1.29), which gives

$$
\begin{equation*}
\left.\mathcal{P}\right|_{z=0}=\frac{1}{2} \mathcal{A}_{0} \operatorname{Re}\left(Z_{\mathrm{rad}}\left(V_{0}\right)^{2} e^{2 i \omega t}\right)+\frac{1}{2} \mathcal{A}_{0} \operatorname{Re}\left(Z_{\mathrm{rad}}\right)|Z|^{2} \tag{9.1.34}
\end{equation*}
$$

For the case of an exponential horn, Eq. (9.1.27) gives

$$
Z_{\text {throat }}=\rho_{0} c \frac{k}{\kappa}= \begin{cases}\rho_{0} c\left[\left(1^{2}-\frac{b^{2}}{4 k^{2}}\right)^{1 / 2}+i \frac{b}{2}\right], & \omega>\omega_{\mathrm{cut}}  \tag{9.1.35}\\ \rho_{0} c i\left[\left(\frac{b^{2}}{4}-k^{2}\right)^{1 / 2}+\frac{b}{2}\right], & \omega<\omega_{\mathrm{cut}}\end{cases}
$$

To gain insight to the design issues pertaining to a horn, let us compare an exponential horn to a hemispherical source in an infinite rigid baffle in the case of a radially symmetric vibration. To make the comparison, we assign the same active surface area to both devices, so $\mathcal{A}_{0}=2 \pi a^{2}$. The time-averaged power radiated by the hemisphere is half the value in Eq. (6.3.16),

$$
\begin{equation*}
\left(\mathcal{P}_{\mathrm{av}}\right)_{\mathrm{hemi}}=\pi \rho_{0} c a^{2}\left|V_{0}^{\prime}\right|^{2} \frac{(k a)^{2}}{(k a)^{2}+1} \tag{9.1.36}
\end{equation*}
$$

Both devices should have a relatively flat response. That is, we wish that $\mathcal{P}_{\text {av }}$ be nearly independent of frequency over as wide a band as possible. The power radiated by an exponential horn is described by Eq. (9.1.32). According to it, if $k / \kappa$ is close to one, then $\mathcal{P}_{\mathrm{av}}$ is independent of $k$. The value of $\kappa$ approaches $k$ as $k$ increases beyond cutoff. Thus, we require that the frequency be substantially greater than the cutoff value, which leads to

$$
\begin{equation*}
\omega \gg \omega_{\text {cutoff }} \Longrightarrow\left(\mathcal{P}_{\mathrm{av}}\right)_{\mathrm{horn}}=2 \pi \rho_{0} c a^{2}\left|V_{0}\right|^{2} \tag{9.1.37}
\end{equation*}
$$

Reference to Fig. 9.5 shows that this criterion is met if $\omega$ is as little as twice the cutoff value.

It is not immediately apparent whether the equivalent hemisphere should be large or small relative to a wavelength, so we shall consider both possibilities. Equation (9.1.36) gives $\mathcal{P}_{\mathrm{av}} \approx \pi \rho_{0} c a^{2}\left|V_{0}^{\prime}\right|^{2}(k a)^{2}$ if $k a \ll 1$. Thus, producing a constant power output over a frequency band with a small hemisphere requires that $\left|V_{0}^{\prime}\right| k a$ be constant. The amplitude of the surface displacement is $\left|V_{0}^{\prime}\right| / \omega$, so sustaining a constant power output with a small diameter requires that the displacement be inversely proportional to the square of the frequency. It is quite challenging to create a small transducer that is capable of increasingly large displacement as the frequency decreases. Hence, let us consider the alternative in which the hemisphere is large relative to a wavelength. If $k a \gg 1$ at the lowest frequency in the band, then the power output would be $\mathcal{P}_{\mathrm{av}} \approx \pi \rho_{0} c a^{2}\left|V_{0}^{\prime}\right|^{2}$. This is a similar dependence to that of a horn, so an invariant $\left|V_{0}\right|$ will yield a constant power output. The difficulty with this design is that it entails a large transducer. Suppose the minimum frequency of interest is 100 Hz . The criterion $k a \gg 1$ in air leads to $a \gg 3.4 \mathrm{~m}$ ! Even if we raise the low end tenfold to 1 kHz , the transducer would be enormous. Large transducers are expensive to manufacture, but there is another fundamental difficulty. The surface
acceleration is proportional to $\omega\left|V_{0}\right|$, so the upper end of the frequency range will be required to undergo relatively high accelerations. Imparting a large acceleration to a large device requires very strong mechanical drivers to induce the vibration, and it also introduces large stresses. In contrast, the cutoff frequency of an exponential horn only sets the exponential constant $b$. We are free to select the horn's length and throat area to fit other specifications, such as a desired sound pressure level at the opening.

The analysis of an exponential horn has been based on the assumption that a backward propagating wave is not present. This is only true if the reflection coefficient at that end is zero. In most cases, the undriven end of the horn is open. The end correction for the open end of a constant cross-sectional waveguide, which approximated the impedance at an open end by the radiation impedance of a piston, may be used for a horn. The impedance at that end is given by Eq. (3.5.53) as $\rho_{0} c \chi(k a)$. This value is very close to the fluid's characteristic impedance, $\rho_{0} c$, if $k a>3$, which corresponds to the radius of the opening exceeding half the acoustic wavelength. In this condition, the reflection coefficient is close to zero. The alternative case in which $k a<3$ requires inclusion of backward waves resulting from reflection at the open end. (This observation may be applied to waveguides whose cross section is not circular by taking the mean radius to be $[\mathcal{A}(L) / \pi]^{1 / 2}$.) The next example describes how a nonzero reflection coefficient at the opening affects the field within the horn.

EXAMPLE 9.1 A horn may be used to obtain a gradual transition between constant cross-sectional waveguides having different sizes. This transitional segment typically is conical to join circular cross sections or a truncated pyramid for square ones. However, we will consider an exponential section. It is desired to determine the relationship between the pressure and particle velocity at the ends $x=0$ and $x=L$. This property is a key part of the algorithm developed in Sect. 3.4.3 to analyze waveguide networks. The requisite quantity is the two-port mobility matrix [ $D$ ], which was defined such that

$$
\left\{\begin{array}{c}
V_{x}(x=0) \\
V_{x}(x=L)
\end{array}\right\}=[D]\left\{\begin{array}{c}
P(x=0) \\
P(x=L)
\end{array}\right\}
$$

Derive an expression for [ $D$ ] in terms of the values of $k, L$, and $b$.

## Significance

The analysis will explore how reflection phenomena are affected by a nonuniform cross section. It has practical application for waveguide networks.

## Solution

We cannot assume that waves propagate in one direction, because that property only occurs if the open end is large and exposed to an open space. Thus, we must retain both waves in Eq. (9.1.22). Manipulations are simpler if we use symbols to represent the characteristic exponents, so we set

$$
\kappa_{1}=\left(k^{2}-\frac{b^{2}}{4}\right)^{1 / 2}-\frac{b}{2} i, \quad \kappa_{2}=\left(k^{2}-\frac{b^{2}}{4}\right)^{1 / 2}+\frac{b}{2} i
$$

The general solution reduces to

$$
P=B_{1} e^{-i \kappa_{1} x}+B_{2} e^{i \kappa_{2} x}
$$

Note that this form is valid for all $k$ values. The particle velocity is obtained from Euler's equation

$$
\rho_{0} c V_{x}=\frac{\kappa_{1}}{k} B_{1} e^{-i \kappa_{1} x}-\frac{\kappa_{2}}{k} B_{2} e^{i \kappa_{2} x}
$$

The pressure and velocity at the ends resulting from these expressions are

$$
\begin{gathered}
B_{1}+B_{2}=P_{0} \\
\frac{\kappa_{1}}{k} B_{1}-\frac{\kappa_{2}}{k} B_{2}=\rho_{0} c V_{0} \\
B_{1} e^{-i \kappa_{1} L}+B_{2} e^{i \kappa_{2} L}=P_{L} \\
\frac{\kappa_{1}}{k} B_{1} e^{-i \kappa_{1} L}-\frac{\kappa_{2}}{k} B_{2} e^{i \kappa_{2} L}=\rho_{0} c V_{L}
\end{gathered}
$$

The process of eliminating $B_{1}$ and $B_{2}$ is facilitated by writing these equations in matrix form

$$
\left\{\begin{array}{l}
P_{0} \\
P_{L}
\end{array}\right\}=\left[T_{p}\right]\left\{\begin{array}{l}
B_{1} \\
B_{2}
\end{array}\right\}, \quad\left\{\begin{array}{l}
\rho_{0} c V_{0} \\
\rho_{0} c V_{L}
\end{array}\right\}=\left[T_{v}\right]\left\{\begin{array}{l}
B_{1} \\
B_{2}
\end{array}\right\}
$$

where the transfer matrices are

$$
\begin{gathered}
{\left[T_{p}\right]=\left[\begin{array}{cc}
1 & 1 \\
e^{-i \kappa_{1} L} & e^{i \kappa_{2} L}
\end{array}\right]} \\
{\left[T_{v}\right]=\left[\begin{array}{cc}
\kappa_{1} / k & -\kappa_{2} / k \\
\left(\kappa_{1} / k\right) e^{-i \kappa_{1} L} & -\left(\kappa_{2} / k\right) e^{i \kappa_{2} L}
\end{array}\right]}
\end{gathered}
$$

The pressure equations are solved for the wave amplitudes, and the result is substituted into the velocity equations. These operations yield

$$
\begin{aligned}
\left\{\begin{array}{l}
B_{1} \\
B_{2}
\end{array}\right\} & =\left[T_{p}\right]^{-1}\left\{\begin{array}{l}
P_{0} \\
P_{L}
\end{array}\right\} \\
\left\{\begin{array}{l}
\rho_{0} c V_{0} \\
\rho_{0} c V_{L}
\end{array}\right\} & =\left[T_{v}\right]\left[T_{p}\right]^{-1}\left\{\begin{array}{c}
P_{0} \\
P_{L}
\end{array}\right\}
\end{aligned}
$$

Hence, the two-port impedance matrix is

$$
[D]=\frac{1}{\rho_{0} c}\left[T_{v}\right]\left[T_{p}\right]^{-1}
$$

The network algorithm is computational, so it would be adequate in most situations to use numerical values of $\left[T_{p}\right]$ and $\left[T_{v}\right]$ corresponding to specified values of $k, b$, and $L$. However, an algebraic evaluation will enable us to examine the effect of the backward-traveling wave. The inverse of $\left[T_{p}\right]$ is

$$
\left[T_{p}\right]^{-1}=\frac{1}{e^{i \kappa_{2} L}-e^{-i \kappa_{1} L}}\left[\begin{array}{cc}
e^{i \kappa_{2} L} & -1 \\
-e^{-i \kappa_{1} L} & 1
\end{array}\right]
$$

In turn, this leads to

$$
[D]=\frac{1}{\rho_{0} c k\left(e^{i \kappa_{2} L}-e^{-i \kappa_{1} L}\right)}\left[\begin{array}{cc}
\left(\kappa_{1} e^{i \kappa_{2} L}+\kappa_{2} e^{-i \kappa_{1} L}\right) & -\left(\kappa_{1}+\kappa_{2}\right) \\
\left(\kappa_{1}+\kappa_{2}\right) e^{-i \kappa_{1} L} e^{i \kappa_{2} L} & -\left(\kappa_{1} e^{-i \kappa_{1} L}+\kappa_{2} e^{i \kappa_{2} L}\right)
\end{array}\right]
$$

This expression is general. To specialize it to the case where $k>b / 2$, we define

$$
\kappa_{1}=K-\frac{b}{2} i, \quad \kappa_{2}=K+\frac{b}{2} i, \quad K=\left(k^{2}-\frac{b^{2}}{4}\right)^{1 / 2}
$$

Application of Euler's identity then reduces the transfer matrix to

$$
\begin{aligned}
{[D]=} & \frac{1}{\rho_{0} c k i \sin (K L)} \\
& {\left[\begin{array}{cc}
{[K \cos (K L)+(b / 2) \sin (K L)]} & -2 K e^{b L / 2} \\
2 K e^{-b L / 2} & {[-K \cos (K L)+(b / 2) \sin (K L)]}
\end{array}\right] }
\end{aligned}
$$

A constant cross section corresponds to $b=0$, which gives $K=k$. In that case, the above expression is identical to Eq. (3.4.14), which is the two-port mobility for a waveguide whose cross section is uniform. An interesting possibility is that it [ $D$ ] is singular if $K L$ is a multiple of $\pi$. This corresponds to a resonance, like that encountered in an impedance tube.

The acoustic mobility matrix has other uses in addition to the stated purpose of analyzing waveguide networks. It describes how the state variables at the ends must be related for any set of end conditions. In the case of an isolated exponential horn, either the complex pressure amplitude or the axial velocity amplitude will be known at the end where the signal is generated. At the far end, a rigid termination requires that $V$ be zero. Another possibility is that the far end is open. In that case, we might set $P=0$, or it might be appropriate to use the open end impedance described by Eq. (3.5.53). Another alternative is that the far end is terminated by a material whose impedance is $Z_{L}$. In any case, one state variable is unknown at each end. The equation $\{V\}=[D]\{P\}$ gives two equations for two pressures and two particle velocities. We may solve those equations for the unknown variables in terms of the excitation amplitude, $P_{0}$ or $V_{0}$ at $x=0$ and $P(x=L)$. Evaluation of such solutions over a frequency band gives a complex frequency response. Depending on the impedance at the far end, there might be frequencies at which the pressure is greatly enhanced. These are cavity resonances, which will be studied from a general framework in the next chapter.

### 9.1.3 Group Velocity

The term "dispersive" refers to waves whose waveform changes shape with increasing propagation distance. In such a wave, the phase speed of a sinusoidal wave will depend on the frequency. The waves at the higher phase speed advance farther than the slower waves, so the arrival times at a specific location differ. The fact that a wave is dispersive has another important effect, which is that the speed at which it transports energy also depends on frequency. The phase speed is merely the speed at which the fact that there is a disturbance is passed to downstream locations. For example, a specific value of the pressure at a certain phase passes from point to point at the phase speed. A simple plane wave is nondispersive, so all features of such a wave, including its energy content, propagate at the speed of sound. The situation is different if a wave is dispersive. The group velocity $c_{\mathrm{g}}$ is the speed at which energy is transported. (The group velocity is a scalar, so it would be more appropriate to say that it is the group speed, but that term is not used.)

The group velocity is important for a variety of waveguide systems, so we shall begin by considering it in general terms. The analysis will lead to a general formula by which the group velocity may be extracted from a dispersion equation. The analysis will close by applying that formula to an exponential horn. As implied by the adjective "group", identification of $c_{\mathrm{g}}$ requires that we consider waves at different frequencies. These frequencies will be taken to be very close, so the associated wave numbers also will be very close. Close frequencies lead to beating. The packet of waves in one beat contains a specific amount of energy that can be tracked.

Two waves of equal amplitude at frequencies $\omega_{1}$ and $\omega_{2}$ propagate in the $x$ direction. We assume that the dispersion equation is known; the function describing the frequency dependence of the wavenumber is $\kappa(\omega)$. Our only interest is how the waves propagate, so it is irrelevant if the field properties are such that the amplitude depends on position transverse to $x$. Hence, the signal we shall consider is

$$
\begin{equation*}
p(x . t)=\operatorname{Re}\left\{B\left[e^{-i \kappa\left(\omega_{1}\right) x+i \omega_{1} t}+e^{-i \kappa\left(\omega_{2}\right) x+i \omega_{2} t}\right]\right\} \tag{9.1.38}
\end{equation*}
$$

The location at which the waves originated also is irrelevant to this investigation, so $B$ may be considered to be real.

Because there is little difference between the frequencies of the individual harmonics, a waveform at any location exhibits a beating pattern as a function of $t$ at any instant. Furthermore, because $\kappa(\omega)$ is taken to be an analytical function, a small difference in frequency leads to wavenumbers that are close. Therefore, spatial profiles also show a beating pattern. The upper graph in Fig. 9.6 shows the spatial profiles at two adjacent instants $t_{1}$ and $t_{1}+\Delta t$, while the lower graph are waveforms at two adjacent positions, $x_{1}$ and $x_{1}$ and $x_{1}+\Delta x$. The dotted lines are the envelope function for each beat pattern. In a time interval $\Delta t$, features of a profile, such as maxima, minima, and zeros, advance at the phase speed. Thus, these features in the upper graph shift to the right by $c_{\text {phase }} \Delta t$. The group velocity is the rate at which the envelope of the beat moves, so the envelope in the upper graph shifts to the right


Fig. 9.6 Addition of two harmonic waves at slightly different frequencies creates a beating pattern in the waveform and spatial profile
by $c_{\mathrm{g}} \Delta t$. If $c_{\text {phase }}>c_{g}$, the extrema will seem to run past the envelope's leading zero, to be replaced by extrema that enter the packet at its trailing zero. In the lower graph, features of the waveform are delayed by the time required for the phase to travel distance $\Delta x$, so this delay is $\Delta x / c_{\text {phase }}$. The delay of the envelope is $\Delta x / c_{g}$. In either view, the total energy density (kinetic and potential) contained between the envelope's zeros is transported at the group velocity.

The first step in identifying the group velocity entails writing the pressure in a form that makes the beating pattern explicit. To that end, the center and difference frequencies are defined to be

$$
\begin{equation*}
\omega_{0}=\frac{1}{2}\left(\omega_{2}+\omega_{1}\right), \quad \Delta \omega=\omega_{2}-\omega_{1} \tag{9.1.39}
\end{equation*}
$$

This corresponds to $\omega_{1}=\omega_{0}-\Delta \omega / 2$ and $\omega_{2}=\omega_{0}+\Delta \omega / 2$. It is assumed that $\Delta \omega \ll \omega_{\mathrm{av}}$, so we may use a two-term Taylor series to approximate the wavenumber at each frequency

$$
\begin{equation*}
\kappa\left(\omega_{1}\right) \approx \kappa_{0}-\kappa_{0}^{\prime} \frac{\Delta \omega}{2}, \quad \kappa\left(\omega_{2}\right) \approx \kappa_{0}+\kappa_{0}^{\prime} \frac{\Delta \omega}{2} \tag{9.1.40}
\end{equation*}
$$

where

$$
\begin{equation*}
\kappa_{0} \equiv \kappa\left(\omega_{0}\right), \quad \kappa_{0}^{\prime}=\left.\frac{d \kappa}{d \omega}\right|_{\omega=\omega_{0}} \tag{9.1.41}
\end{equation*}
$$

Substitution of these representations into Eq. (9.1.38) eventually leads to

$$
\begin{equation*}
p=E(x, t) \operatorname{Re}\left\{2 B\left[e^{-i k_{0} x+i \omega_{0} t}\right]\right\} \tag{9.1.42}
\end{equation*}
$$

The function $E(x, t)$, which is an amplitude modulation function, is given by

$$
\begin{equation*}
E(x, t)=\cos \left[\frac{\Delta \omega}{2}\left(t-\kappa_{0}^{\prime} x\right)\right] \tag{9.1.43}
\end{equation*}
$$

The complex factor in Eq. (9.1.42) represents a nominal constant amplitude wave at the center frequency. This term propagates at the phase speed $c_{\text {phase }}=\omega_{0} / \kappa_{0}$. The envelope of the wave is $\pm E(x, t)$. The frequency of $E(x, t)$ is $\Delta \omega / 2$, which is much slower than $\omega_{0}$, and its wavenumber is $\kappa_{0}^{\prime} \Delta \omega / 2$, which is much smaller than $\kappa_{0}$. Equation (9.1.43) indicates that all features of the envelope, including its zeros and maxima, propagate in the direction of increasing $x$ at speed $1 / \kappa_{0}^{\prime}$. Therefore, the interval between adjacent zeros of $E(x, t)$ constitutes a wave packet that contains a fixed amount of energy. Because $\kappa_{0}^{\prime}$ is the derivative of the $\kappa(\omega)$ dispersion curve at the average frequency, it follows that the phase speed and group velocity at any frequency are

$$
\begin{equation*}
c_{\text {phase }}=\frac{\omega}{\kappa}, \quad c_{\mathrm{g}}=\frac{1}{d \kappa / d \omega} \tag{9.1.44}
\end{equation*}
$$

In principle, $c_{\mathrm{g}}$ could be negative, which would indicate that energy propagates oppositely to the direction of propagation. Such an occurrence is exceptional.

The analysis has taken $\kappa(\omega)$ to be real. However, the function described by Eq. (9.1.20) is complex. This often is the case, as is evidenced by the wavenumbers for plane waves in a dissipative fluid, Sect. 3.3. The real part of the wavenumber is associated with the propagation, and the group velocity also is extracted from $\operatorname{Re}(\kappa)$. Thus, the general relations are

$$
\begin{equation*}
c_{\text {phase }}=\operatorname{Re}\left(\frac{\omega}{\kappa}\right), \quad c_{\mathrm{g}}=\operatorname{Re}\left(\frac{1}{d \kappa / d \omega}\right) \equiv \operatorname{Re}\left(\frac{d \omega}{d \kappa}\right) \tag{9.1.45}
\end{equation*}
$$

The second form in the equation for $c_{\mathrm{g}}$ is useful if it is easier to solve the dispersion equation for the frequency corresponding to a specified wavenumber.

The preceding is generally applicable. In the case of an exponential horn, the dispersion equation is Eq. (9.1.20). It describes $\kappa$ as a function of the plane wave number $k$, whereas Eq. (9.1.45) takes $\kappa$ to be a function of $\omega$. The chain rule for differentiation leads to

$$
c_{\mathrm{g}}=c \operatorname{Re}\left(\frac{1}{d \kappa / d k}\right)=\left\{\begin{array}{l}
c\left(1-\frac{b^{2}}{4 k^{2}}\right)^{1 / 2} \text { if } k>\frac{b}{2}  \tag{9.1.46}\\
0 \text { if } k<\frac{b}{2}
\end{array}\right.
$$

This expression tells us that the group velocity in an exponential horn decreases with decreasing frequency up to cutoff. Below the cutoff frequency, the group velocity is zero. The latter aspect could have been anticipated by the earlier analysis, which showed that the average power flow is zero below cutoff.

A comparison of the preceding relation for group velocity and Eq. (9.1.26) for the phase speed shows that

$$
\begin{equation*}
c_{\mathrm{g}} c_{\text {phase }}=c^{2} \tag{9.1.47}
\end{equation*}
$$

It is not unusual in acoustics to encounter this relation, but it is not a general property. The relation in a sense helps us to resolve any uneasiness regarding the phase speed increasing without bound as $\omega$ decreases the cutoff value. We now see that increasing $c_{\text {phase }}$ is accompanied by decreasing $c_{\mathrm{g}}$, until energy ceases to propagate at $\omega_{\text {cutoff }}$.

EXAMPLE 9.2 The velocity of a source at the throat of an exponential horn in air ( $\rho_{0}=1.2 \mathrm{~kg} / \mathrm{m}^{3}$ and $\left.c=340 \mathrm{~m} / \mathrm{s}\right)$ is $v_{0}=0.004[\sin (500 t)+\cos (501 t)]$ $\mathrm{m} / \mathrm{s}$. The area growth constant is $b=0.6 \mathrm{~m}^{-1}$, and reflections at the open end are negligible. Determine the pressure waveforms at $x=0.60 \mathrm{~m}$ and 1.20 m . Use them to determine the group velocity, and compare the result to the value obtained from the dispersion relation for an exponential horn.

## Significance

The notion here is that seeing how features of a propagating wave are correlated to the group velocity will enhance intuitive understanding of these concepts.

## Solution

The specified analysis calls for extraction of the group velocity directly from computed waveforms, followed by a comparison of the result to the one obtained from the dispersion equation. Thus, our first task is to construct the waveforms. The pressure in an exponential horn is described by Eq. (9.1.22) for a single harmonic. To combine two harmonics, we write the sum in time-domain form. Before we do so, we determine the amplitude coefficient for each harmonic. The pressure in each harmonic is related to the source's velocity by Eq. (9.1.35) for the throat impedance. The values for frequencies of 500 and $501 \mathrm{rad} / \mathrm{s}$ are

$$
Z_{\text {throat }}\left(\omega_{1}=500\right)=399.42+83.23 i, \quad Z_{\text {throat }}\left(\omega_{2}=501\right)=399.45+83.07 i \text { Rayl }
$$

The corresponding pressure amplitudes are $P_{n}(x=0)=Z_{\text {throat }} V_{n}$, where $V_{1}=$ $0.004 / i$ and $V_{2}=0.004 \mathrm{~m} / \mathrm{s}$. The exact occurrence of beats requires that $\left|P_{2}\right|=\left|P_{1}\right|$. The amplitudes obtained here are sufficiently close to treat the data as a true beat.

The waveform must be sampled at a very high rate in order to determine accurately the instants at which $p=0$. The data set we shall use corresponds to a sampling interval of $0.001\left(2 \pi / \omega_{\mathrm{av}}\right)$. The data set that is generated is

$$
p=\operatorname{Re}\left[Z_{\text {throat }}\left(\omega_{1}\right) \frac{0.004}{i} e^{i \omega_{1} t_{n}}+Z_{\text {throat }}\left(\omega_{1}\right)(0.004) e^{i \omega_{1} t_{n}}\right], \quad t_{n}=\frac{0.002 \pi n}{\omega_{\mathrm{av}}}
$$

Figure 1 shows the pressure waveform at $x_{1}=0.6 \mathrm{~m}$. The existence of a beat is evident, but the oscillation at the average frequency is too rapid to see the actual waveform.


Figure 1.
Figure 2 zooms the waveforms at $x_{1}=0.6 \mathrm{~m}$ and $x_{2}=1.2 \mathrm{~m}$ in the interval surrounding the first zero of the envelope. We can evaluate the group velocity by dividing the distance $x_{2}-x_{1}$ by the time required for a zero to travel between these locations. Hence, the feature we seek is a zero of the pressure, but which one? The answer emerges when we consider the first zero after $t=1.57 \mathrm{~s}$ at which $p$ transitions from positive to negative. Both waveforms are such that the smallest peak is to the left and the smallest valley is to the right. This is what we would expect at the location where the amplitude modulation function vanishes. The other zeros correspond to oscillations at the mean frequency, so the amplitude of the oscillation will be larger than it is in the interval where the modulation function is small.


Figure 2.
A less intuitive way in which to identify the instant at which the envelope function is zero is to examine the zeros of the signal. In the interval depicted in Fig. 2, we see that $p\left(x_{1}, t\right)=0$ at $t_{n}=1.5501,1.5564,1.5627,1.5690,1.5730,1.5753$, $1.5815,1.5878,1.5941$, and 1.6004 seconds. The intervals between adjacent zeros are $t_{n+1}-t_{n}=0.0063,0.0063,0.0063,0.0040,0.0023,0.0062,0.0063,0.0063$, and 0.0063 . A regular interval, which is 0.0063 s for the present data, corresponds to the zeros of the high-frequency oscillation at the average frequency. An extra zero due to the envelope function zero breaks up such regularity. This condition is manifested
here by the subintervals of 0.0040 s followed by 0.0023 s . The instant they mark is $t_{1}=1.5730 \mathrm{~s}$. For greater precision, we interpolate between the instants at which the pressure changes sign. The value extracted in this manner is $t_{1}=1.573015 \mathrm{~s}$. A similar analysis of the $x_{2}$ data in the zoomed interval leads to identification of $t_{2}=1.574817 \mathrm{~s}$ as the instant at which the envelope is zero. The group velocity is thereby estimated as

$$
\left(c_{\mathrm{g}}\right)_{\mathrm{data}} \approx \frac{x_{2}-x_{1}}{t_{2}-t_{1}}=330.0 \mathrm{~m} / \mathrm{s}
$$

The analytical expression for the group velocity is Eq. (9.1.46). The value of $k$ is set by the average frequency, so we have

$$
c_{\mathrm{g}}=\left.c\left(1-\frac{b^{2}}{4 k^{2}}\right)^{1 / 2}\right|_{k=\omega_{\mathrm{av}} / c}=332.9 \mathrm{~m} / \mathrm{s}
$$

To further calibrate the accuracy of this procedure, let us evaluate the phase velocity by a similar calculation. The envelope modulates the amplitude of a wave at the center frequency that propagates at the phase speed. The best feature to track is a zero near an instant when the envelope is largest because its location can be identified with little ambiguity. The one we shall use occurs slightly later than 4.5 s . The values identified by scanning the data and interpolating across the zero crossing are $t_{1}=4.500292 \mathrm{~s}$ and $t_{2}=4.502019 \mathrm{~s}$. This gives $c_{\text {phase }}=\left(x_{2}-x_{1}\right) /\left(t_{2}-t_{1}\right)=347.4$ $\mathrm{m} / \mathrm{s}$, whereas Eq. (9.1.20) gives $c_{\text {phase }}=347.3 \mathrm{~m} / \mathrm{s}$.

Our identification scheme performed well. It should be evident from this exercise that experiments to extract phase and group velocity from waveform data measured over a small spatial range require highly accurate and precise equipment. Accuracy would be improved by using measurements at points that are farther apart. However, if that interval exceeds a wavelength at the average frequency, we will not be able to correlate zeros at two locations because of the almost periodic nature of the waveform.

### 9.1.4 WKB Solution for an Arbitrary Horn

There are several reasons to use a horn whose profile is neither conical or exponential. For example, to smoothly join two cylindrical tubes, it would be necessary that the transition segment be circular with a radius function $R_{\text {wall }}(x)$ for which $d R_{\text {wall }} / d x=$ 0 at the ends. Except for a few cases, $\mathcal{A}(x)$ will be such that the Webster horn equation has no simple solution. A technique that has been developed is the WKB method. ${ }^{1}$ It is an approximate analysis in the frequency domain based on reasonable assumptions.

[^30]Our starting point is consideration of the propagation of a plane wave in the direction of increasing $x$. In an ideal unbounded fluid, propagation introduces a phase lag $-k x$ to the pressure. Therefore, it is reasonable to anticipate that the phase lag of the arbitrary plane wave will be negative and increase with increasing $x$. In view of these considerations, it must be that the general time-domain representation of the wave may be written as

$$
\begin{equation*}
p(x, t)=|P(x)| \operatorname{Re}\left(e^{i(\omega t-\Phi(x)}\right) \tag{9.1.48}
\end{equation*}
$$

The phase lag for this representation is the real variable $\Phi(x)$.
The phase speed, $c_{\text {phase }}$, is obtained by equating the phase of $p$ at $(x, t)$ and $\left(x+d x, t+d x / c_{\text {phase }}\right)$. Therefore,

$$
\begin{align*}
\omega t-\Phi(x) & =\omega\left(t+\frac{d x}{c_{\text {phase }}}\right)-\Phi(x+d x) \\
& =\omega t+\frac{\omega d x}{c_{\text {phase }}}-\left[\Phi(x)+\frac{d}{d x} \Phi(x) d x\right] \tag{9.1.49}
\end{align*}
$$

Solution of this relation for the phase speed yields

$$
\begin{equation*}
c_{\text {phase }}=\frac{\omega}{\frac{d}{d x}[\Phi(x)]} \equiv \frac{\omega}{\frac{d}{d x}[\arg (P(x))]} \tag{9.1.50}
\end{equation*}
$$

The local wavenumber therefore is

$$
\begin{equation*}
k_{\text {local }}=\frac{\omega}{c_{\text {phase }}}=\frac{d}{d x}[\arg (P(x))] \tag{9.1.51}
\end{equation*}
$$

These expressions are suitable for any plane wave propagating in a fixed direction. One of the outcomes of the WKB method will be a description of the phase speed and local wavenumber in terms of the frequency and the properties of the waveguide.

The WKB analysis of the Webster horn equation assumes that there is no reflection at the far end, so only waves that propagate in the direction of increasing $x$ are present. The starting point is an ansatz suggested by Eq. (9.1.48), but lacking the distinction between the phase and the magnitude of $P$. Specifically, we begin with

$$
\begin{equation*}
p(x, t)=\operatorname{Re}\left(P(x) e^{i \omega t}\right), \quad P(x)=B e^{-i \Theta(x)} \tag{9.1.52}
\end{equation*}
$$

where $B$ is constant. The function $\Theta(x)$ is allowed to be complex. If we write it as $\operatorname{Re}(\Theta)+i \operatorname{Im}(\Theta)$, then we have

$$
\begin{equation*}
|P(x)|=|B| e^{\operatorname{Im}(\Theta)}, \quad \Phi(x)=\operatorname{Re}(\Theta)-\arg (B) \tag{9.1.53}
\end{equation*}
$$

Substitution of Eq. (9.1.52) into the Webster horn equation gives

$$
\begin{equation*}
\left[-i \frac{d^{2} \Theta}{d x^{2}}-\left(\frac{d \Theta}{d x}\right)^{2}-i \Gamma \frac{d \Theta}{d x}+k^{2}\right] B e^{-i \Theta}=0 \tag{9.1.54}
\end{equation*}
$$

where

$$
\begin{equation*}
\Gamma(x)=\frac{1}{\mathcal{A}} \frac{d \mathcal{A}}{d x} \tag{9.1.55}
\end{equation*}
$$

The exponential factor is not identically zero, so a nontrivial solution corresponds to vanishing of the bracketed term

$$
\begin{equation*}
i \frac{d^{2} \Theta}{d x^{2}}+\left(\frac{d \Theta}{d x}\right)^{2}+i \Gamma(x) \frac{d \Theta}{d x}-k^{2}=0 \tag{9.1.56}
\end{equation*}
$$

This is a nonlinear second-order ordinary differential equation. The frequencydomain Webster horn equation is linear, so it might seem that we have made the analysis more difficult. However, such thinking ignores the difference in the behavior of the complex pressure $P$ and phase variable $\Theta$. In the case of an exponential waveguide, $P$ varies rapidly over the scale of a wavelength, whereas $\Theta=\kappa x$, which is a much more gradual dependence on $x$. In the case of arbitrary $\mathcal{A}(x)$, we can find an exponential area function $B e^{b x}$ that closely fits the actual $\mathcal{A}(x)$ over any small range of $x$ values. Within this range, the exponential horn solution for this approximating $b$ should be close to the solution for the true $\mathcal{A}(x)$. This observation suggests that in the short $x$ interval where the area functions match, it is reasonable to consider $\Theta=\kappa x$. A different exponential factor $b$ will fit $\mathcal{A}(x)$ in an adjacent $x$ interval. Thus, the value of $\kappa$ for this interval will be different. However, unless the cross-sectional area undulates greatly over a very small scale, the adjacent value of $\kappa$ should not differ greatly from the previous value; that is, $d \kappa / d x$ may be expected to be small. In other words, we can anticipate that in general $\Theta \approx \kappa(x) x$, with $\kappa(x)$ being a slowly varying function. Because $d^{2} \Theta / d x^{2}=\left(d^{2} \kappa / d x^{2}\right) x+2 d \kappa / d x$, and $d \kappa / d x$ is taken to be small in some sense, perhaps it is reasonable to consider $d^{2} \Theta / d x^{2}$ to be even smaller. This is the WKB approximation.

The approximation is implemented as an iterative procedure. The first iteration results from setting $d^{2} \Theta / d x^{2}$ to zero in Eq. (9.1.56). We solve that equation for $d \Theta / d x$ and integrate to find the first approximation $\Theta^{(1)}(x)$. A second approximation, $\Theta^{(2)}(x)$, may be obtained by using the first approximation to estimate $d^{2} \Theta / d x^{2}$. The specific equations to be solved are

$$
\begin{gather*}
\left(\frac{d \Theta^{(1)}}{d x}\right)^{2}+i \Gamma(x) \frac{d \Theta^{(1)}}{d x}-k^{2}=0  \tag{9.1.57}\\
\left(\frac{d \Theta^{(2)}}{d x}\right)^{2}+i \Gamma(x) \frac{d \Theta^{(2)}}{d x}-k^{2}=-i \frac{d^{2} \Theta^{(1)}}{d x^{2}}
\end{gather*}
$$

It is not difficult to prove that the equation governing the $n$th iterant resembles the one for $\Theta^{(2)}$. It is

$$
\begin{equation*}
\left(\frac{d \Theta^{(n)}}{d x}\right)^{2}+i \Gamma(x) \frac{d \Theta^{(n)}}{d x}-\left(k^{2}-i \frac{d^{2} \Theta^{(n-1)}}{d x^{2}}\right)=0 ; n=1,2, \ldots \tag{9.1.58}
\end{equation*}
$$

The result of solving this quadratic equation for $d \Theta^{(n)} / d x$ is a recurrence relation

$$
\begin{equation*}
\Theta^{(0)}=0 ; \quad \frac{d \Theta^{(n)}}{d x}=-\frac{i}{2} \Gamma(x) \pm\left[k^{2}-\frac{1}{4} \Gamma(x)^{2}-\frac{d^{2} \Theta^{(n-1)}}{d x^{2}}\right]^{1 / 2}, \quad n=1,2, \ldots \tag{9.1.59}
\end{equation*}
$$

We seek the wave that propagates in the positive $x$ direction, which corresponds to $\operatorname{Re}(\Theta)>0$. Thus, the root for the positive sign is selected. A direct integration, with $\Theta^{(n)}(0)=0$, leads to

$$
\begin{equation*}
\Theta^{(n)}=-\frac{i}{2} \int_{0}^{x} \Gamma(\eta) d \eta+\int_{0}^{x}\left[k^{2}-\frac{1}{4} \Gamma(\eta)^{2}-\frac{d^{2} \Theta^{(n-1)}}{d \eta^{2}}\right]^{1 / 2} d \eta \tag{9.1.60}
\end{equation*}
$$

The definition in Eq. (9.1.55) may be written as $\Gamma=(d / d x) \ln (\mathcal{A})$, which leads to

$$
\begin{equation*}
\Theta^{(n)}=-i \ln \left(\frac{\mathcal{A}(x)^{1 / 2}}{\mathcal{A}(0)^{1 / 2}}\right)+\int_{0}^{x}\left[k^{2}-\frac{1}{4} \Gamma(\eta)^{2}-\frac{d^{2} \Theta^{(n-1)}}{d \eta^{2}}\right]^{1 / 2} d \eta \tag{9.1.61}
\end{equation*}
$$

Let $N$ be the last iteration step. Because $e^{-\ln (u)} \equiv 1 / u$, the complex pressure amplitude resulting from the $n$th iteration is given by

$$
\begin{equation*}
P(x)=B \frac{\mathcal{A}(0)^{1 / 2}}{\mathcal{A}(x)^{1 / 2}} \exp \left(-i \int_{0}^{x}\left[k^{2}-\frac{1}{4} \Gamma(\eta)^{2}-\frac{d^{2} \Theta^{(N-1)}}{d \eta^{2}}\right]^{1 / 2} d \eta\right) \tag{9.1.62}
\end{equation*}
$$

An important aspect of the preceding expression is that the integrand generally will be complex for $N>1$, which is evidenced by Eq. (9.1.61). Consequently, $|P|$ and $\Phi(x)$ in Eq. (9.1.53) can only be extracted after the integral has been evaluated. The exception is $N=1$. In addition, it is evident that any iteration beyond the first will be quite complicated. It is for this reason that a WKB analysis seldom goes beyond the first iteration. For $N=1$, the integrand in Eq. (9.1.62) is real. The corresponding description of the pressure amplitude and phase angle is

$$
\begin{equation*}
|P|=|B| \frac{\mathcal{A}(0)^{1 / 2}}{\mathcal{A}(x)^{1 / 2}}, \quad \Phi(x)=\int_{0}^{x}\left[k^{2}-\frac{1}{4} \Gamma(\eta)^{2}\right]^{1 / 2} d \eta-\arg (B) \tag{9.1.63}
\end{equation*}
$$

The fact that the pressure depends inversely on the square root of the area is consistent with the requirement that the time-averaged power flowing through a cross section be independent of $x$. Substitution of $\Phi(x)$ into Eqs. (9.1.50) and (9.1.51) yields

$$
\begin{align*}
& k_{\text {local }}=\left[k^{2}-\frac{1}{4} \Gamma(x)^{2}\right]^{1 / 2}  \tag{9.1.64}\\
& c_{\text {phase }}=\frac{\omega}{\left[k^{2}-\frac{1}{4} \Gamma(x)^{2}\right]^{1 / 2}}
\end{align*}
$$

In practice, the integral in Eq. (9.1.62) will be quite challenging to evaluate analytically at any iteration, including the first. One exception is the exponential horn, for which $\Gamma(x)=b$. The first iteration gives

$$
\begin{equation*}
\Theta^{(0)}=0 \Longrightarrow \Theta^{(1)}=-\frac{i}{2} b x+\left(k^{2}-\frac{b^{2}}{4}\right)^{1 / 2} x \tag{9.1.65}
\end{equation*}
$$

This gives $d^{2} \Theta^{(1)} / d x^{2}=0$, which means that $\Theta^{(2)}$ will be the same as $\Theta^{(1),}$, and so on. From this, we conclude that $\Theta^{(1)}=\Theta$, but the preceding expression states that $\Theta^{(1)}=\kappa x$, where $\kappa$ is the complex wavenumber in Eq. (9.1.20). Hence, the WKB solution is identically correct for an exponential horn. This is as it should be, because the fact that $\Theta=\kappa x$ for an exponential horn means that $d^{2} \Theta / d x^{2}=0$ is not an approximation.

When the $\mathcal{A}$ function is such that it is not possible to carry out the integral analytically, it could be evaluated numerically. However, it is reasonable to ask why we should employ the WKB method in that case, because an alternative is to use numerical methods to solve the Webster horn equation. Doing so does not entail assuming that $d^{2} \Theta / d x^{2}$ is very small. The next example will examine the various facets of both numerical approaches.

EXAMPLE 9.3 In order to explore the effect of a constriction in a waveguide, consider the case where the cross-sectional area is $\mathcal{A}(x)=$ $\mathcal{A}_{0}\left[\varepsilon+\cos (\pi x / L)^{2}\right]$, which varies periodically over length $L / 2$ between a minimum of $\varepsilon \mathcal{A}_{0}$ and a maximum of $(1+\varepsilon) \mathcal{A}_{0}$. At $x=0$, the particle velocity is $v_{x}=\operatorname{Re}\left(V_{0} e^{i \omega t}\right)$, and the far end, which is $x=2 L$, is terminated with an absorptive material that has a zero reflection coefficient. The derivation of the Webster horn equation requires that the cross section changes gradually. This condition is satisfied if $L$ is much greater than the nominal wavelength $2 \pi / k$, so set $k L=20 \pi$. Use the WKB method to determine the complex pressure amplitude $P(x)$, and the corresponding local wavenumber and phase speed for the case where $\varepsilon=0.1$. Restrict the analysis to the first iteration. Then, assess the quality of the approximate solution by using numerical methods to solve the Webster horn equation.

## Significance

The WKB method is interesting for the insights derived from its solution. However, it is very difficult to perform higher iterations, so its accuracy is uncertain. A direct
numerical solution of the Webster does not entail fundamental approximations, so we will have the opportunity to decide whether the WKB method is worth pursuing.

## Solution

The first iteration of the WKB solution is indicated by Eq. (9.1.62) to be

$$
\begin{equation*}
P(x)=B \frac{\mathcal{A}(0)^{1 / 2}}{\mathcal{A}(x)^{1 / 2}} \exp \left(-i \int_{0}^{x}\left[k^{2}-\frac{1}{4} \Gamma(\eta)^{2}\right]^{1 / 2} d \eta\right) \tag{1}
\end{equation*}
$$

The value of $B$ is set by matching $V_{x}$ at $x=0$ to $V_{0}$. This requires application of Euler's equation. Differentiation of Eq. (1) and application of Leibnitz' rule for differentiation of an integral leads to

$$
\begin{aligned}
\left.\frac{d P}{d x}\right|_{x=0}= & -\left.\frac{1}{2} B \mathcal{A}(0)^{1 / 2}\left(\frac{1}{\mathcal{A}(x)^{3 / 2}} \frac{d \mathcal{A}}{d x}\right)\right|_{x=0} \\
& -i B \frac{\mathcal{A}(0)^{1 / 2}}{\mathcal{A}(x)^{1 / 2}}\left[k^{2}-\frac{1}{4} \Gamma(0)^{2}\right]=-i \omega \rho_{0} V_{0}
\end{aligned}
$$

By definition, $\Gamma(0)=(d \mathcal{A} / d x) / \mathcal{A}$ evaluated at $x=0$, so the preceding yields

$$
B=\frac{\omega \rho_{0} V_{0}}{\left[k^{2}-\frac{1}{4} \Gamma(0)^{2}\right]^{1 / 2}-\frac{i}{2} \Gamma(0)}
$$

The function $\Gamma(x)$ for the given area function is

$$
\begin{equation*}
\Gamma(x)=\frac{1}{\mathcal{A}} \frac{d \mathcal{A}}{d x}=\frac{-2(\pi / L) \varepsilon \sin (2 \pi x / L)}{\varepsilon+\cos (\pi x / L)^{2}} \tag{2}
\end{equation*}
$$

At $x=0$, we have $\Gamma(0)=0$, so we have established that

$$
\begin{equation*}
B=\rho_{0} c V_{0} \tag{3}
\end{equation*}
$$

Substitution of Eq. (2) into (1) leads to an integral that is in the class of elliptic integrals, which are tabulated functions. However, that recognition still requires considerable sophistication to manipulate the integral to a standard form. Since the primary purpose is to determine the accuracy of the WKB method, we shall abandon the effort to evaluate the integral analytically. Instead, we will evaluate the integral numerically.

It is best to use nondimensional variables for numerical studies, so let $x=\tilde{x} L$ and $P=\rho_{0} c V_{0} \tilde{P}$. Then, the first WKB iteration, Eq. (1), becomes

$$
\begin{equation*}
\tilde{P}_{\mathrm{wkb}}=\frac{\tilde{\mathcal{A}}(0)^{1 / 2}}{\tilde{\mathcal{A}}(\tilde{x})^{1 / 2}} e^{-i \tilde{\Phi}(\tilde{x})} \tag{4}
\end{equation*}
$$

where the phase angle is

$$
\begin{equation*}
\tilde{\Phi}(\tilde{x})=\int_{0}^{\tilde{x}}\left[(k L)^{2}-\frac{1}{4} \tilde{\Gamma}(\eta)^{2}\right]^{1 / 2} d \eta \tag{5}
\end{equation*}
$$

The nondimensional area properties are

$$
\begin{aligned}
& \tilde{\mathcal{A}}(\tilde{x})^{1 / 2} \equiv \frac{\mathcal{A}(\tilde{x} L)}{\mathcal{A}_{0}}=\varepsilon+\cos (\pi \tilde{x})^{2} \\
& \tilde{\Gamma}(\tilde{x})=L \Gamma(\tilde{x} L)=\frac{-\pi \varepsilon \sin (2 \pi \tilde{x})}{1+\cos (\pi \tilde{x})^{2}}
\end{aligned}
$$

Evaluation of Eq. (5) is straightforward for a numerical routine such as MATLAB's quadl because the integrand has no singularities. Evaluation at a large number of locations $\tilde{x}_{j}$ may be done efficiently by exploiting the incremental nature of an integral, which gives

$$
\tilde{\Phi}\left(\tilde{x}_{j}\right)=\tilde{\Phi}\left(\tilde{x}_{j-1}\right)+\int_{\tilde{x}_{j-1}}^{\tilde{x}_{j}}\left[(k L)^{2}-\frac{1}{4} \tilde{\Gamma}(\tilde{\eta})^{2}\right]^{1 / 2} d \hat{\eta}
$$

Now let us turn to the direct application of numerical methods. In terms of the nondimensional variables defined above, the Webster horn equation in the frequency domain is

$$
\frac{d^{2} \tilde{P}}{d \tilde{x}^{2}}+\tilde{\Gamma}(\tilde{x}) \frac{d \tilde{P}}{d \tilde{x}}+(k L)^{2} \tilde{P}=0
$$

Standard numerical software offers routines for solving first-order differential equations. To convert this second-order differential equation to first-order form, we define two variables $Y_{1}$ and $Y_{2}$, such that

$$
\begin{equation*}
Y_{1}=\tilde{P}=\frac{P}{\rho_{0} c v_{0}}, \quad Y_{2}=\frac{d \tilde{P}}{d \tilde{x}}=\frac{L}{\rho_{0} c V_{0}} \frac{d P}{d x} \tag{6}
\end{equation*}
$$

The derivative identity is $d Y_{1} / d \tilde{x}=Y_{2}$, and the derivative of $Y_{2}$ is the second derivative of $\tilde{P}$, whose value is given by the differential equation. Thus, the firstorder equations to solve are

$$
\frac{d}{d \tilde{x}}\left\{\begin{array}{l}
Y_{1}  \tag{7}\\
Y_{2}
\end{array}\right\}=\left\{\begin{array}{c}
Y_{2} \\
-(k L)^{2} Y_{1}-\tilde{\Gamma}(\tilde{x}) Y_{2}
\end{array}\right\}
$$

Solution of this pair of order differential equations requires values of $Y_{1}$ and $Y_{2}$ at $\tilde{x}=0$. Euler's equation gives $d P / d x=-i \omega \rho_{0} V_{0}$ at $x=0$. It follows from Eq. (6) that

$$
\begin{equation*}
\left.Y_{2}\right|_{x=0}=-i k L \tag{8}
\end{equation*}
$$

What should we use for the boundary value of $Y_{1}$ ? If we do not set it correctly, we will not initiate a wave that propagates solely in the direction of increasing $x$. To resolve this dilemma, we recall the earlier observation that we may fit an exponential shape to $\mathcal{A}(x)$ in a small interval. Let us do so for the region around $x=0$. Then we can use the equivalent $b$ factor to set the relation between $P$ and $V$ at $x=0$ as though the horn were exponential. For an exponential horn, $\Gamma=b$ everywhere, whereas $\Gamma(0)=0$ in the present system. Therefore, the exponential constant should be $b_{\text {eq }}=0$. According to Eq. (9.1.27) for an exponential horn, the relation between $P$ and $V$ at $x=0$ is

$$
\begin{equation*}
V_{0}=\frac{\kappa_{\mathrm{eq}}}{k} \frac{\left.P\right|_{x=0}}{\rho_{0} c}=\left.\left\{\left[1-\left(\frac{b_{\mathrm{eq}} L}{k L}\right)^{2}\right]^{1 / 2}-\frac{1}{2}\left(\frac{b_{\mathrm{eq}} L}{k L}\right)\right\} V_{0} \tilde{P}\right|_{\tilde{x}=0} \tag{9}
\end{equation*}
$$

Setting $b_{\text {eq }}=0$ leads to

$$
\left.Y_{1}\right|_{\tilde{x}=0}=1
$$

Many techniques have been developed to solve a set of first-order differential equations. A fourth-to-fifth-order Runge-Kutta routine, which is implemented as ode45 in MATLAB, works well for the present configuration. It requires definition of a function that evaluates the right side of the first-order equation. A program loop increments $x$ and calls ode 45 with the function name as the first argument. The value of $Y_{1}$ at each output location $x_{n}$ is saved.

Errors sometimes accumulate in numerical analyses, so it is useful to have ways of checking our work. One way of doing so is to replace the given $\tilde{\Gamma}(\tilde{x})$ with a constant value for an exponential horn. Comparison of the computed pressure to the analytical solution will serve to verify that the program steps are correct. It also is good practice to verify that reducing the $x$ increment in the program loop and decreasing the error tolerances do not significantly alter the computed result. Another check is to compute the time-averaged power flowing through any cross section. From Eq. (6), we have $P=\rho_{0} c V_{0} Y_{1}$ and $d P / d x=\left(\rho_{0} c V_{0} / L\right) Y_{2}$. Then, Euler's equation gives

$$
V_{x}=-\frac{1}{i \omega \rho_{0}} \frac{d P}{d x}=i \frac{V_{0}}{k L} Y_{2}
$$

Correspondingly, the power flowing across any cross section is

$$
\begin{align*}
\mathcal{P}_{\mathrm{av}} & =\frac{1}{2} \mathcal{A}(\tilde{x}) \operatorname{Re}\left(P V_{x}^{*}\right)=\frac{1}{2} \mathcal{A}(\tilde{x}) \operatorname{Re}\left[\left(\rho_{0} c V_{0} Y_{1}\right)\left(V_{0} \frac{i}{k L} Y_{2}\right)^{*}\right]  \tag{10}\\
& =\frac{1}{2} \frac{\rho_{0} c\left|V_{0}\right|^{2}}{k L} \mathcal{A}(\tilde{x}) \operatorname{Re}\left(\frac{Y_{1} Y_{2}^{*}}{i}\right)
\end{align*}
$$

At all locations, the numerical solution gave $\mathcal{P}_{\text {av }} /\left(\rho_{0} c\left|V_{0}\right|^{2} \mathcal{A}_{0}\right)=0.5500 \pm$ $4\left(10^{-13}\right)$.

Figure 1 displays the numerical results for $Y_{1}$ below the $\mathcal{A}$ and $\Gamma$ functions. The pressure rises where $\mathcal{A}$ is relatively small, but its features otherwise are unremarkable.
(

To compare the WKB and direct numerical solutions, we examine $|P|$ and $\tilde{\Phi}$. Equation (4) gives

$$
\left|\tilde{P}_{\mathrm{WKB}}\right|=\frac{\mathcal{A}(0)^{1 / 2}}{\mathcal{A}(x)^{1 / 2}}
$$

and Eq. (5) gives $\Phi_{\text {WKв }}(x)$. These quantities are obtained in the direct numerical solution by post-processing of the $Y_{1}$ data. By definition $Y_{1}=\tilde{P}(x)$. Conversion of the $Y_{1}$ data to polar form places $\Phi$ in a $2 \pi$ range. The consequence is that $\Phi$ can change discontinuously. This data may be converted to a continuous variable by invoking an unwrap routine that is available in many software packages. Alternatively, jumps of $\Phi$ by $2 \pi$ may be canceled manually by comparing the phase angle at each $x$ location to its predecessor at an adjacent location.

Figure 2 compares the amplitude and phase for the alternative methods. The latter is plotted as $\phi / x$, because that quantity would be the constant wavenumber of a simple plane wave. Both analyses are in extraordinarily good agreement, except for the phase near the source.


Figure 2.
It also is useful to compare the local wavenumber and phase speed. These quantities for the first iteration of the WKB method are given by Eq. (9.1.64). To find $k_{\text {local }}$ for the differential equation solution, we use the general relation, Eq. (9.1.51), which requires values of $(d / d x) \arg (P(x)) \equiv d \Phi / d x$. We shall use a central finite difference to approximate Eq. (9.1.51), so that

$$
k_{\mathrm{local}}\left(x_{n}\right)=-\left.\frac{d \Phi}{d x}\right|_{x_{n}} \approx\left(\frac{\Phi\left(x_{n-1}\right)-\Phi\left(x_{n+1}\right)}{\left(x_{n+1}-x_{n-1}\right)}\right)
$$

The phase speed in either case is $c_{\text {phase }}\left(x_{n}\right)=\omega / k_{\text {local }}\left(x_{n}\right)$.
Comparisons of the wavenumber and phase speed are provided by Fig. 3. Here, the agreement is less good. In particular, the differential equation solution indicates that the highest phase speed occurs where the area is smallest. The first WKB iteration misses this feature. However, the numerical solution also has a troubling aspect, specifically, the small-scale fluctuation in $k_{\text {local }}$ and $c_{\text {phase }}$. It is not clear whether this is an artifact or a true property. Examination of the data in the interval $0.5<\tilde{x}<1.5$ indicates that there are 20 ridges. The value of $k L$ is $20 \pi$, which means that there are 20 half-wavelengths in this interval. Furthermore, the curves were found to be unaltered when the error tolerance and time step for the Runge-Kutta solver were decreased, nor were they altered when a backward difference was used to evaluate $k_{\text {local }}$.


Which method is preferable is discretionary in most cases. Once one has programmed either solution, adaptation of that program to analyze other configurations merely requires altering the definition of $\Gamma(x)$. The WKB first iteration is easier to program and requires very little CPU time. However, the differential equation solution also is quite fast because it only requires solving two coupled equations. The main drawback of the WKB method is that its numerical implementation is limited to the first iteration unless we wish to evaluate the second derivative of an iteration's results. Thus, if extreme accuracy is required, especially for the phase speed, then numerical solution of the Webster horn equation is preferable. However, in most situations, we are only interested in the amplitude and phase angle of $P$, and we can accept a fractional dB error. In such situations, the WKB method might be quite acceptable.

### 9.2 Two-Dimensional Waveguides

We now turn our attention to situations where the pressure within a waveguide varies transversely to the propagation. Our starting point is the simplest system, in which the fluid is situated in the region bounded by two infinite planes. It is assumed that
the pressure is invariant in one direction parallel to these planes. Thus, the signal propagates parallel to the boundary, and it varies perpendicularly to the boundaries. Despite the apparent simplicity, the various configurations we shall examine feature most of the phenomena encountered in realistic systems.

### 9.2.1 General Solution

The model system is the fluid contained between two parallel infinite planes. There is no variation in one of the directions parallel to these planes. This is the twodimensional situation depicted in Fig. 9.7, where $x$ measures distance in the direction of propagation and $y$ measures distance from one wall. A harmonic excitation applied along the plane at $x=0$ induces a signal that propagates in the direction of increasing $x$. Our first task is to ascertain the general nature of its dependence on both $x$ and $y$.

Fig. 9.7 A two-dimensional waveguide consisting of the region between parallel infinite planes


We may shortcut the separation of variables formalism by using a complex exponential $e^{-i \kappa x}$ to represent the propagation, where the constant axial wavenumber $\kappa$ must be determined. Hence, we begin with

$$
\begin{equation*}
P=\Phi(y) e^{-i \kappa x} \tag{9.2.1}
\end{equation*}
$$

Note that the above expression is equally valid if $\kappa$ is replaced to $-\kappa$, which would correspond to waves traveling in the negative $x$ direction. This ansatz for $P$ will satisfy the Helmholtz equation if

$$
\begin{equation*}
\frac{d^{2} \Phi}{d y^{2}}+\mu^{2} \Phi=0, \quad \mu^{2}=k^{2}-\kappa^{2} \tag{9.2.2}
\end{equation*}
$$

The $y$-dependence is generally described by

$$
\begin{equation*}
\Phi=B e^{-i \mu y}+C e^{i \mu y} \tag{9.2.3}
\end{equation*}
$$

This form is useful because it will lead us to a fundamental interpretation. However, most individuals prefer trigonometric functions for analysis, so the form we shall use

$$
\begin{equation*}
\Phi=B \cos (\mu y)+C \sin (\mu y) \tag{9.2.4}
\end{equation*}
$$

The coefficients $B$ and $C$ are not the same as they were in Eq. (9.2.3), but this is of no concern because there will be no occasion where both forms are used simultaneously.

Equation (9.2.4) constitutes the general solution. The unknown parameters in it are $B, C$, and $\mu$. These will be set by satisfaction of boundary conditions at $y=0$ and $y=H$. Determination of $\mu$ will then lead to the axial wavenumber, which is given by

$$
\begin{equation*}
\kappa= \pm\left(k^{2}-\mu^{2}\right)^{1 / 2} \tag{9.2.5}
\end{equation*}
$$

The plus/minus sign allows us to select the proper branch cut in situations where $\mu>k$ or $\mu$ is complex. In any case the sign of the square root must be taken to satisfy the Sommerfeld radiation condition for propagation in the direction of increasing $x$, which requires that $\operatorname{Re}(\kappa) \geq 0$ and $\operatorname{Im}(\kappa) \leq 0$.

### 9.2.2 Rigid Walls

For our initial study, we consider the walls to be rigid. Acoustic parlance is to say that this is a hard-walled waveguide. The general solution, Eq. (9.2.4), will satisfy the rigid wall boundary condition at $y=0$ if

$$
\begin{equation*}
\left.\frac{\partial P}{\partial y}\right|_{y=0}=\mu C=0 \tag{9.2.6}
\end{equation*}
$$

It is possible that $\mu=0$, but in that case, $P$ is independent of $y$. This is the case of a plane wave. Thus, we set $C=0$. Then, the rigid condition at $y=H$ requires that

$$
\begin{equation*}
\left.\frac{\partial P}{\partial y}\right|_{y=H}=-\mu B \sin (\mu H)=0 \tag{9.2.7}
\end{equation*}
$$

It cannot be that $B=0$, because that would set $P$ to zero. Thus, it must be that

$$
\begin{equation*}
\sin (\mu H)=0 \Longrightarrow \mu_{n}=\frac{n \pi}{H}, \quad n=0,1,2, \ldots \tag{9.2.8}
\end{equation*}
$$

The description of $P$ in Eq. (9.2.1) at any $n$ is a waveguide mode, and $n$ is the mode number. The modes that arise in structural and mechanical vibration typically are standing waves in all directions, whereas the modes of a waveguide propagate in one direction and form a stationary pattern transversely to the propagation direction. The values of $\mu_{n}$ are the eigenvalues, and the equation whose roots give the eigenval-
ues is the characteristic equation. Somewhat confusingly, $\Phi$, which describes the transverse variation of $P$ corresponding to $\mu_{n}$, is called a transverse mode function, or more concisely, a transverse mode.

The only aspect that we have not found is the coefficient $B$. The solution for $P$ satisfies the Helmholtz equation and rigid wall conditions independently of the value of $B$. Furthermore, any value assigned to it would be compensated when we determine the actual signal generated by a specified excitation. We will soon define a standard definition of $B$, but a common practice for graphing a mode is to set $B$ such that the maximum value of $\Phi_{n}$ is one. Accordingly, we shall temporarily leave this coefficient in algebraic form, so the transverse modes of a hard-walled waveguide are

$$
\begin{equation*}
\Phi_{n}=B_{n} \cos \left(\frac{n \pi y}{H}\right) \tag{9.2.9}
\end{equation*}
$$

The axial wavenumber in the $n$th mode is

$$
\begin{equation*}
\kappa_{n}=\left[k^{2}-\left(\frac{n \pi}{H}\right)^{2}\right]^{1 / 2} \tag{9.2.10}
\end{equation*}
$$

The pressure and particle velocity corresponding to a specific $n$ are

$$
\begin{gather*}
P=B_{n} \cos \left(\frac{n \pi y}{H}\right) e^{-i \kappa_{n} x}, \quad n=0,1,2, \ldots  \tag{9.2.11}\\
\bar{V}=\frac{B_{n}}{\rho_{0} c}\left[\frac{\kappa_{n}}{k} \cos \left(\frac{n \pi y}{H}\right) \bar{e}_{x}-i \frac{n \pi}{k H} \sin \left(\frac{n \pi y}{H}\right) \bar{e}_{y}\right] e^{-i \kappa_{n} x}
\end{gather*}
$$

An important feature is the $n=0$ mode, $\Phi_{0}=B_{0}$, for which $\kappa_{0}=k$. Both $P$ and $\bar{V}$ are independent of $y$ in this mode, and $\bar{V}$ is in the $x$ direction; it is a one-dimensional plane wave.

As the mode number increases, the value of $\kappa_{n}$ decreases, until $n$ is sufficiently large that the value of $\kappa_{n}$ becomes imaginary. The largest $n$ for which $\kappa_{n}$ is real is the cutoff mode number. For the case of rigid walls, it is

$$
\begin{equation*}
N_{\mathrm{cut}}=\text { floor }\left(\frac{k H}{\pi}\right) \tag{9.2.12}
\end{equation*}
$$

Although Eq. (9.2.10) for $n>N_{\text {cut }}$ is satisfied if $\kappa_{n}$ is either a positive or negative imaginary value, only the negative value leads to decay. Specifically, we set

$$
\begin{equation*}
\kappa_{n}=-i\left(\mu_{n}^{2}-k^{2}\right)^{1 / 2} \text { if } n>N_{\text {cut }} \tag{9.2.13}
\end{equation*}
$$

In this case, the $x$ dependence of $P$ and $\bar{V}$ is

$$
\begin{equation*}
e^{-i \kappa_{n} x}=e^{-\left(\mu_{n}^{2}-k^{2}\right)^{1 / 2} x} \tag{9.2.14}
\end{equation*}
$$

These modes evanesce. The decay constant $\left|\kappa_{n}\right|$ for $n>N_{\text {cut }}$ increases with increasing $n$ when $k$ is fixed. This means that the cutoff modes become less significant at a specific $x$ as the number of "wiggles" of the mode in the $y$ direction increases. The terminology we will use is to say that a propagating mode has an axial wavenumber that is positive real (for propagation in the sense of increasing $x$ ). An evanescent mode has an axial wavenumber that is negative imaginary. (These definitions will be refined when we examine situations where the axial wavenumber is complex.) The value of $\kappa_{n}$ may be computed in one operation regardless of the frequency according to $\kappa_{n}=\operatorname{conj}\left(\left(k^{2}-\mu_{n}^{2}\right)^{1 / 2}\right)$ or $\kappa_{n}=-\operatorname{conj}\left(\left(k^{2}-\mu_{n}^{2}\right)^{1 / 2}\right)$. Which alternative should be selected depends on whether the software gives $\operatorname{Im}(\kappa)<0$ if $\kappa^{2}<0$.

If we consider a specific mode, that is, hold $n$ fixed, $\left|\kappa_{n}\right|$ decreases as $k$ decreases from a large value. The cutoff frequency of a mode is the value below which that mode becomes evanescent,

$$
\begin{equation*}
\omega_{\mathrm{cut}}=\frac{n \pi c}{H} \tag{9.2.15}
\end{equation*}
$$

Rather than viewing the behavior of a specific mode as the frequency decreases, we may consider $N_{\text {cut }}=$ floor $(\omega H /(\pi c))$ as the highest mode number at which a mode at a specific frequency will propagate. The $n=0$ (plane wave) mode never is cutoff.

The phase speed of a propagating mode is

$$
\begin{equation*}
c_{\text {phase }}=\frac{\omega}{\kappa_{n}}=\frac{c}{\left[1-\left(\frac{n \pi}{k H}\right)^{2}\right]^{1 / 2}} \tag{9.2.16}
\end{equation*}
$$

The group velocity obtained from Eq. (9.2.10) is

$$
\begin{equation*}
\left(c_{\mathrm{g}}\right)_{n}=c \frac{1}{d \kappa_{n} / d k}=c\left[1-\left(\frac{n \pi}{k H}\right)^{2}\right]^{1 / 2} \tag{9.2.17}
\end{equation*}
$$

For frequencies below the cutoff, the axial wavenumber is imaginary, so the phase and group velocities are zero. An interesting aspect is the similarity of these properties to those of an exponential horn. In both systems

$$
\begin{equation*}
c_{\mathrm{g}} c_{\text {phase }}=c^{2} \tag{9.2.18}
\end{equation*}
$$

The next section will explain this relation in terms of the trace velocity.

### 9.2.3 Interpretation

The description of a mode in Eq. (9.2.11) indicates that a constant value of $P$, whose amplitude is proportional to $\cos (n \pi y / H)$, propagates in the $x$ direction. This
amplitude is sustained as the signal moves downstream. Thus, it is appropriate to say that a waveguide mode is a nonuniform plane wave. This means that the rays of constant phase are parallel and that the amplitude is not constant along the wavefronts, which are perpendicular to the rays.

An alternate picture explains several aspects, including the transition to evanescent behavior. To obtain it, we employ Euler's identity to convert the transverse mode functions to complex exponentials

$$
\begin{equation*}
\Phi_{n}=\frac{B_{n}}{2}\left[e^{(-i n \pi y / H)}+e^{(i n \pi y / H)}\right] \tag{9.2.19}
\end{equation*}
$$

This is the incarnation of the general solution, Eq. (9.2.3), for the rigid wall case. When we use this representation to form the $n$th waveguide mode, we find that

$$
\begin{equation*}
P_{n}=\frac{B_{n}}{2}\left[e^{-i \bar{k}_{1} \cdot \bar{x}}+e^{-i \bar{k}_{2} \cdot \bar{x}}\right] \tag{9.2.20}
\end{equation*}
$$

where the position is $\bar{x}=x \bar{e}_{x}+y \bar{e}_{y}$. The wavenumber vectors are

$$
\begin{align*}
& \bar{k}_{1}=\kappa_{n} \bar{e}_{x}+\frac{n \pi}{H} \bar{e}_{y} \\
& \bar{k}_{2}=\kappa_{n} \bar{e}_{x}-\frac{n \pi}{H} \bar{e}_{y} \tag{9.2.21}
\end{align*}
$$

These vectors are depicted in Fig.9.8. The vectors are directed at angle $\psi$ above and below the $x$-axis. In view of Eq. (9.2.10), $\left|\bar{k}_{1}\right|=\left|\bar{k}_{2}\right|=k$. Consequently, each wavenumber vector is the product of $k$ and a unit vector, which is the way each is depicted in the figure.

The construction in Fig. 9.8 leads to an alternate description of the wavenumber vectors as

Fig. 9.8 Wavenumber vectors for the plane waves that are equivalent to mode $n$ in a two-dimensional hard-walled waveguide


$$
\begin{gather*}
\bar{k}_{1}=k \bar{e}_{1}, \quad \bar{k}_{2}=k \bar{e}_{2} \\
\bar{e}_{1}=\left(\cos \psi_{n}\right) \bar{e}_{x}+\left(\sin \psi_{n}\right) \bar{e}_{y}, \quad \bar{e}_{2}=\left(\cos \psi_{n}\right) \bar{e}_{x}-\left(\sin \psi_{n}\right) \bar{e}_{y}  \tag{9.2.22}\\
\psi_{n}=\tan ^{-1}\left(\frac{n \pi}{\kappa_{n} H}\right)=\cos ^{-1}\left(\frac{\kappa_{n}}{k}\right)=\sin ^{-1}\left(\frac{n \pi}{k H}\right)
\end{gather*}
$$

Substitution of this representation of the wave numbers into Eq. (9.2.20) leads to

$$
\begin{equation*}
P_{n}=P_{n}^{(1)}+P_{n}^{(2)}, \quad P_{n}^{(j)}=\frac{B_{n}}{2} e^{-i k \bar{e}_{j} \cdot \bar{x}} \tag{9.2.23}
\end{equation*}
$$

This representation describes a mode $P_{n}(x, y)$ as a superposition of simple plane waves whose phase speed is $c$ that propagate obliquely in opposite sense relative to the $x$-axis. The particle velocity is the superposition of the contribution of the individual waves.

$$
\begin{equation*}
\rho_{0} c \bar{V}=\frac{B_{n}}{2}\left(\bar{e}_{1} e^{-i \bar{k}_{1} \cdot \bar{x}}+\bar{e}_{2} e^{-i \bar{k}_{2} \cdot \bar{x}}\right) \tag{9.2.24}
\end{equation*}
$$

The picture of the signal as superposed plane waves explains many features. The transition to an evanescent mode is a consequence of the monotonic increase of $\psi_{n}$ as $n$ increases at fixed $k$. Ultimately, if $n>N_{\text {cut }}$, the last of Eq. (9.2.22) states that $\psi_{n}$ is an angle whose sine is greater than one. This leads to a complex $\psi_{n}$, with $\cos \psi_{n}$ being a negative imaginary value. The corresponding $x$ components of $\bar{k}_{1}$ and $\bar{k}_{2}$ become negative imaginary values, which converts the mode from one that propagates to one that evanesces.

Decomposition of a propagating mode into simple plane waves gives a new perspective for the phase speed in Eq. (9.2.16) and the group velocity in Eq. (9.2.17). For a propagating mode, $\cos \psi_{n}=\kappa_{n} / k=\left[1-(n \pi)^{2} /(k H)^{2}\right]^{1 / 2}$. Thus, these propagation speeds may be written as

$$
\begin{equation*}
c_{\text {phase }}=\frac{c}{\cos \psi_{n}}, \quad c_{\mathrm{g}}=c \cos \psi_{n} \tag{9.2.25}
\end{equation*}
$$

The phase speed is recognizable as the trace in the $x$ direction of the phase velocity of the oblique plane waves. In contrast, the group velocity is the component in the $x$ direction of the phase velocity of each oblique plane wave. Increasing $\psi_{n}$, either by raising $n$ or decreasing $k$, increases the trace velocity along the $x$-axis. Ultimately, at the cutoff frequency, the phase speed is infinite because the wavefronts of the simple waves are parallel to the $x$-axis, so all points on the $x$-axis simultaneously see a specific wavefront. The explanation of the group velocity lies in the fact that the oblique plane waves are nondispersive, so both their group and phase velocities are $c$. Thus, a packet of oblique waves formed from two slightly different frequencies would propagate obliquely at speed $c$. The projection onto the $x$-axis of the packet's location would move in the $x$ direction at speed $c \cos \psi_{n}$.

Representation of a mode as two plane waves makes is possible to view the propagation in terms of rays. Figure 9.9a follows the ray of each plane wave that passes through an arbitrary point $C$. The $\bar{e}_{1}$ ray is incident at the upper wall, where it

Fig. 9.9 Multiple reflections in a hard-walled two-dimensional waveguide observed by decomposing a mode into its constituent plane waves

is reflected, thereby generating an $\bar{e}_{2}$ ray. Because the boundary is rigid, the reflection coefficient is $R=1$, so the pressure in the reflected $\bar{e}_{2}$ ray matches the pressure in the incident ray. In other words, the reflected wave seems to have originated from the image $C_{2}$ of point $C$. A similar process applies to the $\bar{e}_{2}$ ray that intersects point $C$, so the pressure on the $\bar{e}_{1}$ ray at point $D$ seems to have originated from image point $C_{1}$. In other words, the signal within a two-dimensional waveguide with rigid walls is a process of repeated reflections.

The standing wave pattern in the transverse $y$ direction results from the requirement that the field be consistent for reflection from either wall. From the viewpoint of the lower wall $\bar{e}_{2}$ is the propagation direction for the incident wave, and $\bar{e}_{1}$ is the propagation direction for the reflected wave. The viewpoint of the upper wall is opposite, with the incident direction being $\bar{e}_{1}$ and the reflected direction being $\bar{e}_{2}$. The reflection coefficient at each wall is $R=1$ because the walls are rigid. Suppose the angle of incidence (and reflection) is an arbitrary value $\chi$. Then, Eq. (5.2.16) states that the pressure field due to reflection at the lower wall is

$$
\begin{align*}
p(x, y, t) & =\operatorname{Re}\left[\frac{B_{n}}{2} e^{i\left(\omega t-k x \sin \psi_{1}\right)}\left(e^{i k y \cos \chi}+e^{-i k y \cos \chi}\right)\right] \\
& =\operatorname{Re}\left[B_{n} e^{i(\omega t-k x \sin \chi)} \cos (k y \cos \chi)\right] \tag{9.2.26}
\end{align*}
$$

Reflection from the upper wall may be described similarly if we replace $y$ with $y^{\prime}=H-y$, so that $y^{\prime}=0$ at the upper wall and $y^{\prime}=H$ at the lower wall. Then, the pressure field from this viewpoint is

$$
\begin{align*}
p(x, y, t)= & \operatorname{Re}\left[B_{n} e^{i(\omega t-k x \sin \chi)} \cos (k(H-y) \cos \chi)\right] \\
\equiv & \operatorname{Re}\left\{B_{n} e^{i(\omega t-k x \sin \chi)}[\cos (k H \cos \chi) \cos (k y \cos \chi)\right. \\
& +\sin (k H \cos \chi) \sin (k y \cos \chi)]\} \tag{9.2.27}
\end{align*}
$$

These alternate descriptions must represent the same pressure field, so it must be that $\sin (k H \cos \chi)=0$. In other words, $k H \cos \chi=n \pi$. The angle $\chi$ was defined to be the angle of incidence, so the angle $\psi$ for the decomposition into oblique waves is $\pi / 2-\chi$. It follows that

$$
\begin{equation*}
\sin \psi=\frac{n \pi}{k H} \tag{9.2.28}
\end{equation*}
$$

This is identical to the third relation for $\psi_{n}$ in Eq. (9.2.22). Thus, we have shown that the waveguide mode is the superposition of a pair of waves that propagate obliquely to the $x$-axis. The propagation angle for each oblique wave is such that the interference patterns resulting from reflection at the lower and upper walls are compatible.

### 9.2.4 Flexible Walls

The model of a waveguide with rigid walls is a useful idealization. However, the walls in some applications may be quite flexible, and they might be composed of an absorptive material, as is the case for turbine engine exhaust systems. Some coastal regions of the ocean may be considered to be a constant-depth waveguide with a pressure-release upper wall (the surface) and a compliant lower wall (the ocean bottom). In the situation to be addressed here, the planar boundary at $y=0$ has local impedance $\rho_{0} c \zeta_{0}$ and the boundary at $y=H$ has local impedance $\rho_{0} c \zeta_{H}$; both parameters are considered to be the same at all locations on the respective surfaces. The normal directions into the fluid are $\bar{n}=\bar{e}_{y}$ and $\bar{n}=-\bar{e}_{y}$ at $y=0$ and $y=H$, respectively. The local impedance condition is $P=\rho_{0} c \zeta(-\bar{n} \cdot \bar{V})$. We use Euler's equation to describe $\bar{V}$, which leads to the boundary conditions at the walls being

$$
\begin{gather*}
P=-\rho_{0} c \zeta_{0} V_{y}=\frac{\zeta_{0}}{i k} \frac{\partial P}{\partial y} \text { at } y=0  \tag{9.2.29}\\
P=\rho_{0} c \zeta_{H} V_{y}=-\frac{\zeta_{H}}{i k} \frac{\partial P}{\partial y} \text { at } y=H
\end{gather*}
$$

The general solution for the pressure in any two-dimensional waveguide is given by Eq. (9.2.4). We substitute this representation into the boundary conditions. The requirement that the result be satisfied for all $x$ allows cancelation of the $e^{-i \kappa x}$ factor. The result is a pair of linear algebraic equations for $B$ and $C$,

$$
\begin{gather*}
\zeta_{0} \mu C-i k B=0 \\
\zeta_{H} \mu[B \sin (\mu H)-C \cos (\mu H)]-i k[B \cos (\mu H)+C \sin (\mu H)]=0 \tag{9.2.30}
\end{gather*}
$$

The matrix form of these equations is

$$
\begin{gather*}
{\left[\begin{array}{ll}
D
\end{array}\right]\left[\begin{array}{ll}
B & C
\end{array}\right]^{\mathrm{T}}=\left[\begin{array}{ll}
0 & 0
\end{array}\right]^{\mathrm{T}}} \\
{[D]=\left[\begin{array}{cc}
-i k & \zeta_{0} \mu \\
{\left[\zeta_{H} \mu \sin (\mu H)-i k \cos (\mu H)\right]-\left[\begin{array}{l}
\zeta_{H} \mu \cos (\mu H)+i k \sin (\mu H)
\end{array}\right]}
\end{array}\right]} \tag{9.2.31}
\end{gather*}
$$

The coefficient equations are homogeneous, so their solution would give $B=$ $C=0$. We seek a nontrivial solution, which can only exist of the equations are rank-deficient. This condition is marked by the coefficient matrix having a zero determinant. The condition that $|[D]|=0$ constitutes the characteristic equation

$$
\begin{equation*}
|[D]|=-\left(k^{2}+\zeta_{0} \zeta_{H} \mu^{2}\right) \sin (\mu H)+i\left(\zeta_{0}+\zeta_{H}\right) k \mu \cos (\mu H)=0 \tag{9.2.32}
\end{equation*}
$$

Multiplication of this equation by $H^{2}$ reveals that it contains only nondimensional parameters: $\mu H, k H, \zeta_{0}$, and $\zeta_{H}$. The frequency is specified, so the only unknown in the characteristic equation is $\mu H$. The presence of sinusoidal terms whose argument is $\mu H$ tells us that there are an infinite number of nondimensional roots, whose values we designate as $\eta_{n}$; the corresponding eigenvalues are $\mu_{n}=\eta_{n} / H$ and $n$ is the mode number.

The preceding characteristic equation has a root at $\mu=0$, but that root is extraneous and must be disregarded if either $\zeta_{0}$ or $\zeta_{H}$ is finite. This is so because $\mu=0$ corresponds to a simple plane wave. The particle velocity in a plane wave is solely in the $x$ direction. The pressure in such a wave would cause the wall to move in the $y$ direction, which is incompatible with the velocity field. The only exception is the case where both walls are rigid. It is noteworthy that the plane wave mode does not have a cutoff condition. The consequence is that only if both walls are rigid can sound propagate at any frequency, including values that are nearly zero.

The $B$ and $C$ coefficients corresponding to an eigenvalue must satisfy Eq. (9.2.31). Because this pair of equations are rank-deficient when $\mu$ equals any $\mu_{n}$, only one of the scalar equations is independent. This means that we only can solve those equations for one coefficient as a proportionality to the other. We take $B_{n}$ to be the arbitrary coefficient, where " $n$ " subscript allows for a value that varies from mode to mode. When we replace $\mu$ with $\eta_{n} / H$ in the first of Eq. (9.2.30) and solve for $C / B$, we find that the waveguide mode is

$$
\begin{gather*}
P=\Phi_{n}(y) e^{-i \kappa_{n} x}  \tag{9.2.33}\\
\Phi_{n}=B_{n}\left[k H \sin \left(\eta_{n} \frac{y}{H}\right)-i \zeta_{0} \eta_{n} \cos \left(\eta_{n} \frac{y}{H}\right)\right] \\
-\left(k^{2} H^{2}+\zeta_{0} \zeta_{H} \eta_{n}^{2}\right) \sin \left(\eta_{n}\right)+i\left(\zeta_{0}+\zeta_{H}\right) k H \eta_{n} \cos \left(\eta_{n}\right) \\
\kappa_{n}= \pm\left(k^{2}-\eta_{n}^{2} / H^{2}\right)^{1 / 2}
\end{gather*}
$$

As was true for the case where both walls are rigid, the alternative sign for $\kappa_{n}$ corresponds to waves that propagate in the $x$ direction in either sense. If either wall's impedance has a resistive part, the root $\eta_{n}$ will be complex. In such case, care must be
taken to select the alternative sign. Stated differently, we must select the appropriate branch cut for the square root. Let us focus on the wave that propagates in the sense of increasing $x$, which means that we only are interested in $\eta_{n}$ values that lead to $\operatorname{Re}\left(\kappa_{n}\right) \geq 0$. Furthermore, the radiation condition requires that the magnitude of the pressure not grow in the direction of propagation, so it must be that $\operatorname{Im}\left(\kappa_{n}\right) \leq 0$.

To identify the restriction this requirement imposes on the axial wavenumber, let us write $\kappa_{n}$ in polar form as $\kappa_{n}=\left|\kappa_{n}\right| e^{i \phi_{n}}$. Thus, we seek the conditions for which a root $\mu_{n}$ will lead to $-\pi / 2 \leq \phi_{n} \leq 0$. From the definition of $\mu_{n}$, we have $\mu_{n}^{2}=k^{2}-\left|\kappa_{n}\right|^{2} \exp \left(2 i \phi_{n}\right) \equiv k^{2}-\left|\kappa_{n}\right|^{2} \cos \left(2 \phi_{n}\right)-i\left|\kappa_{n}\right|^{2} \sin \left(2 \phi_{n}\right)$. If $\phi_{n}$ is in the proper range, then $-1 \leq \sin 2 \phi_{n} \leq 0$, so $\operatorname{Im}\left(\mu_{n}^{2}\right) \geq 0$. There is no limitation to $\left|\kappa_{n}\right|$, so the preceding tells us that $\mu_{n}^{2}$ is in either the first or second quadrant, that is, $\mu_{n}^{2}=u \exp (i \theta)$ with $u>0$ and $0 \leq \theta \leq \pi$. DeMoivre's theorem gives two values: $\mu_{n}=u^{1 / 2} \exp (i \theta / 2)$ and $\mu_{n}=-u^{1 / 2} \exp (i \theta / 2)$. The second root, being the negative of the first, is accounted for as the second complex exponential in Eq. (9.2.3). Therefore, only the first value is retained. Because $\theta$ must lie in the first or second quadrant of the complex plane, it follows that the roots $\mu_{n}$ of the characteristic equation that are consistent with the radiation condition lie in the first quadrant of the complex plane. Roots that lie in other quadrants are extraneous.

A complex wavenumber $\kappa_{n}$ leads to an oscillating function that decays as it propagates in the $x$ direction. To understand the nature of the transverse dependence when $\kappa_{n}$ is complex, we recall identities that convert sinusoidal functions having a complex argument to hyperbolic functions whose argument is real. They give

$$
\begin{align*}
& \sin (\alpha+\iota \beta) \equiv \sin (\alpha) \cosh (\beta)+i \cos (\alpha) \sinh \beta \\
& \cos (\alpha+\iota \beta) \equiv \cos (\alpha) \cosh (\beta)-i \sin (\alpha) \sinh \beta \tag{9.2.34}
\end{align*}
$$

If neither $\alpha$ nor $\beta$ is negative, then both hyperbolic functions increase with increasing $\beta$. Thus, setting $\alpha+i \beta=\mu_{n} y$ shows that $\Phi_{n}$ varies in an oscillatory manner with a growing amplitude.

Cases where the eigenvalues are real are exceptional. One consists of the combination of a rigid and a pressure-release wall, which we take respectively to be $y=0$ and $y=H$. The general characteristic equation, Eq. (9.2.32), is divided by $\zeta_{0}$ to handle the infinite value of that impedance. Because $\zeta_{H}=0$, what remains is $k \mu \cos (\mu H)=0$. The cosine function must vanish because $\mu=0$ is extraneous. Thus, the roots $\eta_{n}$ are odd multiples of $\pi / 2$. The modal properties are

$$
\begin{gather*}
\mu_{n}=\left(\frac{2 n-1}{2}\right) \frac{\pi}{H}, \quad n=1,2, \ldots, \quad \kappa_{n}=\left[k^{2}-n^{2}\left(\frac{\pi}{H}\right)^{2}\right]^{1 / 2} \\
\Phi_{n}=B_{n} \cos \left(\frac{(2 n-1)}{2} \frac{\pi y}{H}\right)  \tag{9.2.35}\\
N_{\mathrm{cut}}=\text { floor }\left(\frac{k H}{\pi}+\frac{1}{2}\right), \quad \omega_{\mathrm{cut}}=\left(\frac{2 n-1}{2}\right) \frac{\pi c}{H}
\end{gather*}
$$

As was true in the case where both walls are rigid, the variation in the $y$ direction is a real harmonic function, which means that the pressure at all points situated at the same $x$ oscillates in-phase.

Whereas a water channel might be approximated as having a lower rigid wall and an upper pressure-release wall, a waveguide with two pressure-release walls has no common incarnation. Such a system could be approximated by using lightly tensioned thin mylar sheets for the walls. However, this model is primarily useful as a limiting behavior. If $\zeta_{0}=\zeta_{H}=0$, then the general characteristic equation reduces to $\sin (\mu H)=0$. This is the same equation as that for a rigid-rigid waveguide, except that $\mu_{n}$ cannot be zero here, so

$$
\begin{gather*}
\mu_{n}=n \frac{\pi}{H}, \quad \kappa_{n}=\left[k^{2}-n^{2}\left(\frac{\pi}{H}\right)^{2}\right]^{1 / 2}, \quad n=1,2, \ldots \\
\Phi_{n}=B_{n} \sin \left(\frac{n \pi y}{H}\right)  \tag{9.2.36}\\
N_{\mathrm{cut}}=\text { floor }\left(\frac{k H}{\pi}\right), \quad \omega_{\mathrm{cut}}=n \frac{\pi c}{H}
\end{gather*}
$$

The case where one or both walls are purely reactive is representative of some common systems. For simplicity, let us take the wall at $y=0$ to be rigid. The specific impedance of the wall at $y=H$ is taken to be $\zeta_{H}=i \sigma$, where $\sigma$ is a real value that may be positive or negative. The characteristic equation in this case is found by dividing Eq. (9.2.32) by $\zeta_{0}$ and then substituting for $\zeta_{H}$. It is

$$
\begin{equation*}
-\sigma \mu H \sin (\mu H)+k H \cos (\mu H)=0 \tag{9.2.37}
\end{equation*}
$$

It is more difficult to extract the roots $\mu_{n} H=\eta_{n}$ of this equation, but once we do the other features follow. The transverse mode function, axial wavenumber, and cutoff parameters are described by

$$
\begin{gather*}
\Phi_{n}=B_{b} \cos \left(\frac{\eta_{n} y}{H}\right) \\
\kappa_{n}=\left(k^{2}-\frac{\eta_{n}^{2}}{H^{2}}\right)^{1 / 2}  \tag{9.2.38}\\
\eta_{n}<k H \text { if } n \leq N_{\mathrm{cut}}, \quad \omega_{\mathrm{cut}}=\eta_{n} \frac{c}{H}
\end{gather*}
$$

Due to the transcendental nature of the characteristic equation, numerical methods are required to find the eigenvalues. Those methods work best when they are initiated with a good estimate. It also is useful to have a general understanding of the nature of the root relative to the stiffness and the frequency. Both objectives are met by rewriting the characteristic equation as

$$
\begin{equation*}
\tan (\mu H)=\left(\frac{k H}{\sigma}\right) \frac{1}{\mu H} \tag{9.2.39}
\end{equation*}
$$

The sole nondimensional parameter affecting the roots is the ratio of $k H$ to $\sigma$. The case where $\sigma \ll k H$ corresponds to either a very small $\sigma$, meaning a wall that is very flexible, or else a very high frequency. The smaller roots $\mu H=\eta_{n}$ will be such that $\mu H \ll k H / \sigma$, so the characteristic equation in this case is well approximated as $\cot (\mu H)=0$. The roots are the same as those of $\cos (\mu H)=0$ for a pressure-release wall at $y=H$. The opposite case is that in which $\sigma \gg k H$, which requires that the wall be quite stiff, as well as that the frequency not be very high. In that case, the characteristic equation is approximately $\tan (\mu H)=0$, and the approximation improves with increasing mode number because of the presence of $\mu H$ in the denominator. The roots in this case are the same as those of $\sin (\mu H)=0$ for a hard-walled waveguide.

These qualitative trends may be identified by a graphical construction of the functions of $\mu H$ on either side of Eq. (9.2.39). Figure 9.10 shows the plots for a positive and a negative value of $\sigma$. The value of $k H /|\sigma|$ is $2.5 \pi$, which is not very large. Nevertheless, the first root is close to the first singularity of $\tan (\mu H)$, that is, $\cos \left(\mu_{1} H\right) \approx 0$, which is the first root for a rigid-pressure-release configuration. Increasing the value of $\mu H$ brings $(k H / \sigma) /(\mu H)$ closer to the zero axis, so the roots progressively become closer to the zeros of $\tan (\mu H)$. Thus, they approach the zeros of $\sin (\mu H)$, which is the characteristic equation for a hard-walled waveguide. Decreasing $\sigma$ moves both hyperbolic curves farther from the zero axis. Nevertheless, the higher roots always tend to approach the zeros of $\tan (\mu H)$. In other words, regardless of how flexible a wall is, the higher eigenvalues always tend to the values for a hard-walled waveguide.


Fig. 9.10 Identification of the real eigenvalues as intersections of the functions forming the characteristic equation for a two-dimensional waveguide bounded by a rigid wall and a locally reactive wall, $k H=\pi$

An unexpected possibility is that an eigenvalue is purely imaginary, because an imaginary $\zeta_{H}$ can lead to a real characteristic equation. To explore this possibility, we set $\mu=i \beta$. Identities give $\cos (\mu H) \equiv \cosh (\beta H)$ and $\sin (\mu H) \equiv i \sinh (\beta H)$. A rearrangement of the terms in the characteristic equation gives

$$
\begin{equation*}
\tanh (\beta H)=-\left(\frac{k H}{\sigma}\right) \frac{1}{\beta H} \tag{9.2.40}
\end{equation*}
$$

Only positive values of $\beta$ are meaningful because of the earlier analysis, which showed that $\mu$ must be in the first quadrant of the complex plane. Consequently, $\tanh (\beta H)>0$, which means that a real root for $\beta H$ can exist only if the function on the right side of the above equation is positive. In turn, this requires that $\sigma<0$, which is the case of a wall that is a pure compliance. The plot of both sides of this equation in Fig. 9.11 confirms this interpretation. It also shows that if $\sigma<0$. there is only one root. We will designate this root as $\beta H=\eta_{\mathrm{im}}$, so that $\mu_{\mathrm{im}}=i \eta_{\mathrm{im}} / H$. In general, we will use a subscript "im" to denote this type of mode. As a generality, this type of mode is not possible if a wall is an inertance.


Fig. 9.11 Identification of the imaginary eigenvalue for a two-dimensional waveguide bounded by a rigid wall and a locally reactive wall, $k H=\pi$

The transverse mode function corresponding to a real eigenvalue is given Eq. (9.2.38). Substitution of $\eta_{n}=i \beta_{\mathrm{im}}$ yields

$$
\begin{equation*}
\Phi_{\mathrm{im}}=B \cosh \left(\eta_{\mathrm{im}} \frac{y}{H}\right) \tag{9.2.41}
\end{equation*}
$$

This function, like those for the real eigenvalues, is real. This means the pressure in any mode oscillates either in-phase or $180^{\circ}$ out-of-phase along any cross section at fixed $x$. Fig. 9.12 shows a few transverse mode functions. Figure 9.11 indicates that the value of the root $\eta_{\text {im }}$ will be very large, unless $k H / \sigma$ is very small. The consequence of a large value of $\eta_{\text {im }}$ is that $\Phi_{\text {im }}$ increases greatly in the vicinity of the compliant wall. Another interesting aspect is that the axial wavenumber for this mode is greater than $k$ because $\kappa_{\mathrm{im}}=\left(k^{2}-\mu_{\mathrm{im}}^{2}\right)^{1 / 2}=\left(k^{2}+\eta_{\mathrm{im}}^{2} / H^{2}\right)^{1 / 2}$. Thus, this type of mode propagates much more slowly than the speed of sound.

Let us compare the results for cases where the wall impedance is infinite (rigid), zero (pressure-release), or imaginary (reactive), to the general situation when the real part of the impedance at either wall is nonzero. The former are idealized cases in which dissipation is absent, whereas energy is absorbed in the wall when $\operatorname{Re}(\zeta)>0$. The presence of dissipation complicates the analysis by requiring the solution of a complex characteristic equation. The eigenvalues $\mu_{n}$ are complex, which means that


Fig. 9.12 Mode functions for a two-dimensional waveguide bounded by a rigid wall and a compliant wall, $k H=\pi$
the pressure as a function of $y$ at fixed $x$ shows phase differences. The axial wavenumber $\kappa_{n}$ also is complex, which means that a wave is attenuated as it propagates. We would recognize the attenuation effect as a consequence of dissipation, even if we did not perform an analysis. Phase differences are a more subtle effect of dissipation.

EXAMPLE 9.4 The lower wall of a two-dimensional waveguide is rigid, and the upper wall is purely resistive with $Z=0.1 \rho_{0} c$. Determine the phase and group velocities for the first three modes in the frequency range $0<k H<10$. Also, plot the transverse mode functions at $k H=0.1,1.0$, and 10 .

## Significance

A highlight is determination of the eigenvalues when the characteristic equation is complex. Another new feature is a general approach for extracting group velocities from data. The results will display an interesting phenomenon exhibited by the group velocity.

## Solution

Before we can solve the characteristic equation we must consider what is meant by the specification "first three modes". If the eigenvalues were real, we would identify the mode number by sequencing $\mu_{n}$ in increasing order. A logical extension for complex eigenvalues is to sequence them in ascending order of their real part, $\operatorname{Re}\left(\mu_{1}\right)<\operatorname{Re}\left(\mu_{2}\right)<\ldots$. The characteristic equation we shall solve is the limiting form that results from multiplying Eq. (9.2.32) by $H / \zeta_{0}$ and then taking the limit as $\zeta_{0} \rightarrow \infty$. The result is

$$
f(\mu H)=-\zeta_{H}(\mu H) \sin (\mu H)+i(k H) \cos (\mu H)=0
$$

where $\zeta_{H}=0.1$ in the current case.

The standard methods for solving a single nonlinear equation $f(x)=0$ when $x$ and $f$ are scalars also are valid for complex variables. Many software packages contain the necessary modifications. MATLAB (as of the 2009) does not. The results presented here were obtained by programming a standard Newton-Raphson procedure, which states that the $j$ th estimate for a root is given by

$$
\mu_{n}^{(j)} H=\mu_{n}^{(j-1)} H-\left.\frac{f(\mu H)}{\frac{d}{d \mu} f(\mu H)}\right|_{\mu H=\mu^{(j-1)} H}
$$

where

$$
\frac{d}{d(\mu H)} f(\mu H)=-\left(\zeta_{H}+i k h\right) \sin (\mu H)-\zeta_{h}(\mu H) \cos (\mu H)
$$

Nonlinear equation solvers generally require initial estimates of the roots. A robust approach that is suitable if it is only necessary to find the eigenvalues at a few values of $k H$ entails inspection of data contours. The characteristic equation is a complexvalued function of the complex variable $\mu H$. When it and $\mu H$ are decomposed into real and imaginary parts, we may graph $\operatorname{Re}(f(\mu H))$ and $\operatorname{Im}(f(\mu H))$ as surfaces above the complex plane $\mu H$ whose orthogonal axes are $\operatorname{Re}(\mu H)$ and $\operatorname{Im}(\mu H)$. If both parts of $f$ are zero at the same point in the complex $\mu H$-plane, then that point marks a root of the characteristic equation. The difficulty is that depiction of these surfaces and their identification of the zeros would be challenging.

An alternative approach recognizes that if both parts of $f(\mu H)$ are zero at a root, then it must be that $|f(\mu H)|=0$ at such locations. A plot of $|f(\mu H)|$ constitutes a single surface above the $\mu H$ complex plane. Its value may be depicted by a contour plot. Because $|f(\mu H)|$ cannot be negative, a root must occur at a location where $|f(\mu H)|$ is a minimum. Minima in a contour plot appear as a point interior to a contour that forms a closed curve. However, closed curves also surround maxima, and a minimum value might not be zero. Thus a reasonable procedure is to draw the contours, evaluate $|f(\mu H)|$ at a point inside a closed loop, and test if that value is close to zero. If so, that value of $\mu H$ is a candidate as a starting value for the root search. Recall that only the contours lying in the first quadrant should be examined because the meaningful roots are those that give $\operatorname{Re}\left(\mu_{n}\right) \geq 0$ and $\operatorname{Im}\left(\mu_{n}\right) \geq 0$.

This procedure is illustrated in Fig. 1, for which $k H=\pi$ and $\zeta_{H}=0.1$. Because only the first quadrant is depicted there, a loop that actually is closed might appear to be open if it is situated near the real or imaginary axis. In view of this possibility, Fig. 1 suggests four roots, $\mu_{n} H \approx(2 n-1) \pi / 2+0.1 n i$. The value of $|f(\mu H)|$ at each of these points is less than 0.1 , whereas $|f(\pi+i)|=6.4$, which suggests that the minima might be zeros. The actual roots were found to be $\mu_{1} H / \pi=0.49949+$ $0.0159 i, \mu_{2} H=1.4984+0.0481 i, \mu_{3} H=2.4972+0.0812 i$.


Figure 1.

A different approach followed by some individuals finds estimates by considering an ideal system, that is, one that is purely reactive. For example, in the present situation the value of $\zeta_{H}$ is much less than one. This suggests that the $\mu_{n}$ values are close to the values for $\zeta_{H}=0$, which is the pressure-release condition described by Eq. (9.2.35). The real part of the estimates identified from the contour plot are these values. The potential difficulty with this approach is that it does not address the case where there is no comparable ideal system, which would be the case if $\operatorname{Re}\left(\zeta_{H}\right)$ is not small.

The present task calls for evaluation of the eigenvalues for a range of frequencies. This suggests yet another approach for generating the initial guesses. The eigenvalues are continuous analytical functions of $k H$. Thus, if we increment $k H$ by a small amount, we can use the values identified for the previous frequency as starting estimates for the new frequency. A benefit of following this approach is that it only requires a contour map to identify the eigenvalues at the lowest $k H$. We begin with a contour plot of $|f(\mu H)|$ for $k H=0.01 \pi$, which is Fig. 2. It suggests that the smallest eigenvalue is close to the origin, while the higher eigenvalues are multiples of $\pi$ plus a small imaginary part. Thus, the starting estimates used to initialize the equation solver at $k H=0.01 \pi$ are $\mu_{1} H \approx 0.1 \pi+0.2 i, \mu_{2} H \approx \pi+0.1 i$, $\mu_{3} H=2 \pi+0.05 i$.


The procedure that was followed set a loop in which kH was increased gradually. At each frequency, a loop over the mode number $n$ initialized the $\mu_{n} H$ value with the value found at the prior frequency. This value was input to the routine that solves the
characteristic equation. The output was the value of $\mu_{n} H$ at the current frequency. After the eigenvalues were determined, the $\kappa_{n} H$ were found from the definition, Eq. (9.2.33), after which the frequency was incremented and the procedure repeated. The result is described in Fig. 3. It will be noted that $\mu_{n} H$ is in the first quadrant, and $k_{n} H$ is in the fourth quadrant, as required. The fact that all plots are continuous curves is an indication that the procedure was carried out correctly.


Figure 3.

Each mode shows a transition in its $\kappa_{n}$ value as the frequency increases. Below a certain frequency, $\operatorname{Re}\left(\kappa_{n}\right)$ is small but nonzero, and $\operatorname{Im}\left(\kappa_{n}\right)$ is relatively large negatively. Above that frequency, $\operatorname{Re}\left(\kappa_{n}\right)$ begins to grow and $\operatorname{Im}\left(\kappa_{n}\right)$ is small negatively. This behavior is analogous to cutoff for situations in which the wall does not dissipate energy.

Division of a $\kappa_{n} H$ value by the corresponding value of $k H$ gives $c_{\text {phase }} / c$. Evaluation of the group velocity requires a little more effort. The nondimensional version of Eq. (9.1.45) is

$$
\frac{c_{\mathrm{g}}}{c}=\operatorname{Re}\left(\frac{d(k H)}{d(\kappa H)}\right)
$$

Although we do not have a functional dispersion equation, we can use finite differences to differentiate the data. A central difference formula based on a uniform frequency increment $\Delta=(k H)_{j}-(k H)_{j-1}$ is

$$
\left(\frac{c_{\mathrm{g}}}{c}\right)_{j} \approx \frac{\Delta}{\operatorname{Re}(\kappa H)_{j+1}-\operatorname{Re}(\kappa H)_{j-1}}
$$

The results of these evaluations appear in Fig. 4.


Figure 4.

As the frequency increases, $c_{\text {phase }}$ approaches $c$ from above, and $c_{\mathrm{g}}$ approaches $c$ from below. The large values of $c_{\text {phase }}$ occur in the low-frequency range where $\kappa_{n}$ is essentially a negative imaginary value. This means that the waveguide mode is evanescent. The consequence is that neither the high phase speed nor the negative group velocity is meaningful, because the wave decays over a very short distance.

The transverse mode functions are computed by substituting into Eq. (9.2.33) the values of $\mu_{1} H, \mu_{2} H$, and $\mu_{3} H$ at $k H=0.01,1$, and 10 . The amplitude factor $B_{n}$ for each is arbitrary in regard to a pictorial representation. Because the eigenvalues are complex, setting all $B_{n}=1$ would lead to mode functions that have substantially different magnitudes. To facilitate comparisons, the transverse mode functions in Fig. 5 have been scaled by evaluating each as a function of $y / H$ for $B_{n}=1$ and then dividing each by its maximum magnitude. (The ordinate axis in Fig. 5 is $y / H$, because the transverse direction runs vertically in the configuration depicted in Fig. 9.7.)


Figure 5.
A few features are worth noting. The highest frequency, $k H=10$, exceeds the cutoff value for each of the three modes. The consequence is that the imaginary part of each mode function is much smaller than its real part. Indeed, the real parts are nearly proportional to $\cos (\pi x /(2 H)), \cos (3 \pi x /(2 H))$, and $\cos (5 \pi x /(2 H))$, which are the first three modes for a rigid/pressure-release configuration. This behavior indicates that the impedance is sufficiently low that the surface effectively is soft at high frequencies. If $k H$ is below the cutoff frequency for a mode, the imaginary part of the mode is comparable to the real part, and it is oscillatory in a pattern that resembles the real part

If we were to compare the present results to those for $\zeta_{H}=0$, we would see that the overall effect of dissipation associated with power transfer across the locally reacting wall is quite small. This is especially true for modes that are not cutoff. For those modes, the wave is gradually attenuated as it propagates. The modes whose cutoff frequency exceeds the oscillation rate seem to be much affected by dissipation, as suggested by a negative group velocity. However, these modes are rapidly attenuated, so it would be difficult to observe such features.

EXAMPLE 9.5 The walls of a two-dimensional waveguide are identical elastic plates whose properties are Young's modulus $E$, density $\rho_{p}$ and thickness $h$. The fluid is water, and the material properties of the plate are $E=100 \rho_{0} c^{2}$, $\rho_{p} / \rho_{0}=5$, and $h / H=0.02$. Frequencies of interest are $k H=0.1,1$, and 10 . For each frequency, determine the eigenvalues of the transverse mode function in the range $\mu_{n} H<10$. Also determine the corresponding values of $\kappa_{n} H$ and the ratio of the plate displacement amplitude to the pressure amplitude in each mode.

## Significance

The interaction between different media usually leads to interesting phenomena. In a sense, this is an extension of the analysis of coincidence frequency for an elastic plate in Sect.5.5.2. The application of symmetry properties to simplify an analysis, which is a general tool, plays a prominent role here.

## Solution

The walls are flexible, but they are not locally reacting. Rather, application of a force results in a displacement wave. The consequence it that it is necessary to formulate the equation of motion for the plates. The pressure exerted by the fluid acts as a load. In addition, enforcement of continuity of particle velocity at the walls must be satisfied.

All features are the same above and below the central plane. A general property of a system that has a plane of symmetry is that the response is either symmetric or antisymmetric with respect to that plane. To exploit that feature the origin is placed midway between the plates so that $x z$ is the plane of symmetry. Figure 1 shows the set up. The normal displacement of the plate $w$ is defined to be positive into the fluid.


Figure 1.

The frequency-domain representation of the pressure and displacement in a mode are

$$
\begin{equation*}
p(x, y, t)=\operatorname{Re}\left(P(x, y) e^{i \omega t}\right), \quad w(x, t)=\operatorname{Re}\left(W(x) e^{i \omega t}\right) \tag{1}
\end{equation*}
$$

Every point $(x, y)$ above the bisecting $x z$-plane is matched by a point at $(x,-y)$. A symmetric response has the property that $P(x,-y)=P(x, y)$. In other words $P$ is an even function of $y$. An antisymmetric response is such that $P(x,-y)=$ $-P(x, y)$, which means that it is an odd function of $y$. Because $P$ is an analytic function, it must be that $\partial P / \partial y=0$ at $y=0$ for a symmetric response and $P=0$ at $y=0$ for an antisymmetric response. These, respectively, are the boundary conditions for a rigid and pressure-release wall. Thus, we will consider two configurations of a two-dimensional waveguide whose width is $H / 2$ : rigid at $y=0$ and an elastic plate at $y=H / 2$ for symmetric modes, and pressure-release at $y=0$ and an elastic plate at $y=H / 2$ for antisymmetric modes. We shall denote which type is under consideration with a superscript " $s$ " or " $a$ ", so the effective wall conditions are

$$
\begin{align*}
& \frac{\partial P^{(s)}}{\partial y}=0 \text { at } y=0, \quad \bar{n} \cdot \nabla P^{(s)} \equiv-\frac{\partial P^{(s)}}{\partial y}=-i \rho_{0} \omega(-i \omega W) \text { at } y=H / 2 \\
& P^{(a)}=0 \text { at } y=0, \quad \bar{n} \cdot \nabla P^{(a)} \equiv-\frac{\partial P^{(a)}}{\partial y}=-i \rho_{0} \omega(-i \omega W) \text { at } y=H / 2 \tag{2}
\end{align*}
$$

The pressure must be a solution of the Helmholtz equation. Propagation in the axial direction is represented by $\exp (-i \kappa x)$. For the $y$ direction we can immediately take care of the boundary condition at $y=0$ in Eq. (2) by recognizing that a cosine function gives a zero gradient, whereas a sine function gives a zero value. We leave the argument of each as $\mu y$, so the pressure may be written as

$$
\begin{gather*}
P^{(s)}=B^{(s)} \cos \left(\mu^{(s)} y\right) e^{-i \kappa^{(s)} x}  \tag{3}\\
P^{(a)}=B^{(a)} \sin b\left(\mu^{(a)} y\right) e^{-i \kappa^{(a)} x}
\end{gather*}
$$

These forms satisfy the Helmholtz equation if

$$
\begin{equation*}
\left(\kappa^{(s)}\right)^{2}+\left(\mu^{(s)}\right)^{2}=k^{2}, \quad\left(\kappa^{(a)}\right)^{2}+\left(\mu^{(a)}\right)^{2}=k^{2,} \tag{4}
\end{equation*}
$$

The time-domain equation of motion for the plate is Eq. (5.5.27). The pressures $p_{1}$ and $p_{2}$ appearing there are applied on either side of the plate. In the present situation we shall ignore the fluid loading outside the waveguide. (This is acceptable if there is a liquid within the waveguide and a gas outside, but it might not be a valid assumption if the fluids are similar.) The displacement is defined to be positive if it is into the waveguide, and the pressure pushes the plate away from the fluid. Hence a negative $p$ should cause a positive value of $\partial^{2} w / \partial t^{2}$, which means the pressure term in the equation of motion that is preceded by the negative sign is the one to use. In the two-dimensional problem, there is no variation in the $z$ direction. Thus, the equation of motion is

$$
\rho_{p} h \frac{\partial^{2} w}{\partial t^{2}}+D \frac{\partial^{4} w}{\partial x^{4}}=-\left.p\right|_{y=H / 2}
$$

where $p$ may be either the symmetric or antisymmetric variable. The parameter $D$ is the flexural rigidity. If the plate is composed of a homogeneous material, rather than a composite, then

$$
D=\frac{E h^{3}}{12\left(1-\nu^{2}\right)}
$$

Equation (1) gives the frequency-domain representation of $w$. To use it we must recognize that the amplitude function $W(x)$ must match the $x$ dependence of the pressure, because it is necessary to match $W$ and $P$ at the plate's surface. Thus, $W$ must be representable as

$$
\begin{equation*}
W(x)=\tilde{W} e^{-i \kappa x} \tag{5}
\end{equation*}
$$

where $\kappa$ and $\tilde{W}$ may be the value for symmetric or antisymmetric modes. Substitution of the representations of $W$ and $P$ into the plate equation of motion leads to

$$
\begin{align*}
& \left(D \kappa^{4}-\rho_{p} h \omega^{2}\right) \tilde{W}^{(s)}=-B^{(s)} \cos \left(\mu^{(s)} \frac{H}{2}\right) \\
& \left(D \kappa^{4}-\rho_{p} h \omega^{2}\right) \tilde{W}^{(s)}=-B^{(a)} \sin \left(\mu^{(a)} \frac{H}{2}\right) \tag{6}
\end{align*}
$$

The last set of equations result from enforcing the continuity conditions in Eq. (2). These equations are

$$
\begin{gather*}
B^{(s)} \mu^{(s)} \sin \left(\mu^{(s)} \frac{H}{2}\right)=\rho_{0} \omega^{2} \tilde{W}^{(s)} \\
-B^{(a)} \mu^{(a)} \cos \left(\mu^{(a)} \frac{H}{2}\right)=\rho_{0} \omega^{2} \tilde{W}^{(s)} \tag{7}
\end{gather*}
$$

Equations (6) and (7) describe the coupling of the acoustic and structural responses.
Let us begin with the analysis of the symmetric modes. The matrix form of Eqs. (7) and (6) for $B^{(s)}$ and $\tilde{W}$ is

$$
\left[D^{(s)}\right]\left\{\begin{array}{c}
B^{(s)}  \tag{8}\\
\tilde{W}^{(s)}
\end{array}\right\}=\left\{\begin{array}{l}
0 \\
0
\end{array}\right\}
$$

where

$$
\left[D^{(s)}\right]=\left[\begin{array}{cc}
\mu^{(s)} \sin \left(\mu^{(s)} \frac{H}{2}\right) & -\rho_{0} \omega^{2} \\
\cos \left(\mu^{(s)} \frac{H}{2}\right) & \left(D \kappa^{4}-\rho_{p} \omega^{2}\right)
\end{array}\right]
$$

Equation (8) leads to $B^{(s)}=\tilde{W}^{(s)}=0$ unless the pair of scalar equations is rankdeficient. Thus, we set $\left|\left[D^{(s)}\right]\right|=0$, which gives

$$
\left(D \kappa^{4}-\rho_{p} h \omega^{2}\right) \mu^{(s)} \sin \left(\mu^{(s)} \frac{H}{2}\right)+\rho_{0} \omega^{2} \cos \left(\mu^{(s)} \frac{H}{2}\right)=0
$$

Because the frequency is specified, the value of $\kappa^{2}$ is defined in terms of $\mu^{(s)}$. This means that the characteristic equation for fixed system parameters and frequency is a function of $\mu^{(s)}$. In a nondimensional sense it is a function of $\eta=\mu^{(s)} H$, such that

$$
\begin{equation*}
F^{(s)}(\eta)=\left\{S \frac{\left[(k H)^{2}-\eta^{2}\right]^{2}}{(k H)^{2}}-1\right\} \eta \sin \left(\frac{\eta}{2}\right)+\frac{H}{h} \frac{\rho_{0}}{\rho_{p}} \cos \left(\frac{\eta}{2}\right)=0 \tag{9}
\end{equation*}
$$

The parameter $S$ is a measure of the dynamic stiffness of the plate. Based on an assumption that Poisson's ratio is $\nu=0.3$, it is

$$
S \equiv \frac{D}{\rho_{p} h H^{2} c^{2}}=\frac{1}{12\left(1-\nu^{2}\right)} \frac{E}{\rho_{0} c^{2}}\left(\frac{\rho_{0}}{\rho_{p}}\right)\left(\frac{h}{H}\right)^{2}=0.080
$$

The derivation of the equations for the antisymmetric modes follows a similar path. When the pressure is $P^{(a)}$ in Eq. (3), the coefficient equations are like Eq. (8),

$$
\begin{gather*}
{\left[D^{(a)}\right]\left\{\begin{array}{c}
B^{(a)} \\
\tilde{W}^{(a)}
\end{array}\right\}=\left\{\begin{array}{l}
0 \\
0
\end{array}\right\}} \\
{\left[D^{(a)}\right]=\left[\begin{array}{cc}
\mu^{(a)} \cos \left(\mu^{(a)} \frac{H}{2}\right) & \rho_{0} \omega^{2} \\
\sin \left(\mu^{(s)} \frac{H}{2}\right) & \left(D \kappa^{4}-\rho_{p} \omega^{2}\right)
\end{array}\right]} \tag{10}
\end{gather*}
$$

The characteristic equation for the antisymmetric modes is

$$
\begin{equation*}
F^{(a)}(\eta)=\left\{S \frac{\left[(k H)^{2}-\eta^{2}\right]^{2}}{(k H)^{2}}-1\right\} \eta \cos \left(\frac{\eta}{2}\right)-\frac{H}{h} \frac{\rho_{0}}{\rho_{p}} \sin \left(\frac{\eta}{2}\right)=0 \tag{11}
\end{equation*}
$$

The rank reduction of the Eqs. (8) and (10) at an eigenvalue means that only one equation is independent. We solve the first for the plate displacement in terms of the pressure amplitude, which gives

$$
\begin{gather*}
\frac{W_{n}^{(s)}}{H}=\frac{\mu_{n}^{(s)} H \sin \left(\mu^{(s)} \frac{H}{2}\right)}{(k H)^{2}}\left(\frac{B^{(s)}}{\rho_{0} c^{2}}\right) \\
\frac{W_{n}^{(a)}}{H}=-\frac{\mu_{n}^{(s)} H \cos \left(\mu^{(a)} \frac{H}{2}\right)}{(k H)^{2}}\left(\frac{B^{(a)}}{\rho_{0} c^{2}}\right) \tag{12}
\end{gather*}
$$

When the characteristic equations, Eqs. (9) and (11), are evaluated at a fixed $k H$, they become functions of a single variable, $\eta$. These equations may be solved numerically, but doing so requires initial guesses. We shall obtain these by plotting $F^{(a)}(\eta)$ and $F^{(s)}(\eta)$ as functions of $\eta$. The eigenvalues are the values of $\eta$ at which a function crosses the zero axis. Figures 2, 3, and 4 show such plots for $k H=0.1$, $k H=1$, and $k H=10$. The eigenvalues are marked as circles. The curves marked as $F_{\mathrm{ac}}^{(s)}$ and $F_{\mathrm{ac}}^{(a)}$ denote the second term in each characteristic equation. They are the terms that remain as $\rho_{p} h /\left(\rho_{0} H\right) \rightarrow 0$. Hence, they represent the purely acoustic effect if the plate did not offer any opposition to an applied pressure, that is, if the walls were pressure-release. The eigenvalues in that case would be $\mu_{n} H=(2 n-1) \pi$ for the symmetric modes and $\mu_{n} H=2 n \pi$ for the antisymmetric modes. The contrary case is $\rho_{0} H /\left(\rho_{p} h\right) \rightarrow 0$. This corresponds to a plate that is extremely massive, and
therefore immobile. In such situations the first part of each characteristic equation is dominant. The same behavior is obtained in the limit as $k H \rightarrow 0$, that is, low frequencies. The corresponding eigenvalues are $\mu_{n} H=2 n \pi$ for the symmetric modes and $\mu_{n} H=(2 n-1) \pi$ for the antisymmetric modes. For the most part, the roots in Fig. 2 match those for the rigid wall limit. The exception are the first two antisymmetric modes. We will see that the axial wavenumber for these modes match the wavenumber at which a flexural wave at $k H=0.1$ can propagate along the plate in a vacuum. The plate cannot sustain an external load (the pressure) in such conditions.


Figure 2.


Figure 3.
At the higher frequency, $k H=10$, described in Fig. 4, neither the pressurerelease nor rigid limits for the wall match the computed eigenvalues. Furthermore, the increment between adjacent eigenvalues is irregular. Both features indicate that the structural dynamics behavior of the wall has an important influence on the acoustic response. The modal properties obtained from numerical solution of both characteristic equations are tabulated below. Perhaps the most obvious feature is the irregularity of the sequence in which symmetric and antisymmetric modes occur at low frequencies. Another manifestation of the structural dynamic effect is the fact that the wall displacement to pressure ratio is much lower at the higher frequency. An important
property is that all $\kappa_{n} H$ values are imaginary for $k H=0.1$ and 1.0. This means that all of the low-frequency modes are evanescent, so a signal will not propagate in the low-frequency range.


Figure 4.

| $k H$ | Mode \#n | $\mu_{n} H$ | $\kappa_{n} H$ | $\frac{1}{W_{n} / H}$ |
| :--- | :--- | :--- | :--- | :--- |
|  |  |  |  | $B_{n} /\left(\rho_{0} c^{2}\right)$ |
| 0.10 | Antisymmetric 1 | 1.7770 | -1.7742 i | -112.054 |
|  | Antisymmetric 2 | 3.0445 | -3.0428 i | -14.780 |
|  | Symmetric 1 | 6.2806 | -6.2798 i | 0.8044 |
|  | Antisymmetric 3 | 9.4244 | -9.4239 i | 0.1585 |
| 1.0 | Symmetric 1 | 2.7933 | -2.6082 i | 2.7511 |
|  | Symmetric 2 | 4.3647 | -4.2486 i | 3.5736 |
|  | Symmetric 3 | 5.8555 | -5.7695 i | 1.2426 |
|  | Antisymmetric 1 | 9.3891 | -9.3357 i | 0.1673 |
| 10.0 | Symmetric 1 | 2.9461 | 9.5562 | 0.02932 |
|  | Antisymmetric 1 | 5.6095 | 8.2785 | 0.05294 |
|  | Symmetric 2 | 8.1489 | 5.7961 | -0.06546 |

A graph showing how the eigenvalues depend on $k H$ would be helpful for a closer examination. Such a picture could be obtained by following the preceding analysis for a sequence of closely space $k H$ values. Fortunately, a more efficient method is available. It is based on considering the characteristic equations, Eqs. (9) and (11), to be functions of $k H$, as well as $\mu H$. The value of each function may be considered to form a surface above the $\kappa H, \mu H$-plane. The eigenvalues are the $\kappa H, \mu H$ where the elevation of this surface is zero. Thus, the numerical procedure entails defining a grid of $k H$ and $\mu H$ values, evaluating $F^{(s)}$ and $F^{(a)}$ at each grid point, and then finding the zero contour of each set of function data. The contours of equal values of $F^{(s)}$ and $F^{(a)}$ are computed. The only contours of interest are those for which $F^{(s)}(\eta)=0$ or $F^{(a)}(\eta)=0$. In MATLAB, the operations would be hold on; $k H=0: d \_k: k H \_m a x ; ~ e t a=0: d \_e: e t a \_m a x ; ~$ C_sym = contour(kH, eta/pi, F_val_sym.', [0 0]); C_anti = contour (kH, eta/pi, F_val_anti.', [0 0]);. In these operations, the contour command graphs the data, the last argument being [ 000 ] specifies that only the zero contour should be plotted, and hold on at the beginning
of the procedure retains the first set of contours when the second set are plotted. A useful aspect is that the $k H, \mu H$ coordinates of the points on each contour are embedded in C_sym and C_anti. This data may be used for further study, for example, to determine the group velocity.

Figure 5 shows the resulting contours. When viewed as plots of $\mu_{n} H$ as a function of $k H$, each curve is multivalued over a small range of $k H$. The irregular spacing of the tabulated data stems from this property. The shading marks the region in which $\mu H>k H$, or equivalently, where $\kappa H$ is imaginary. The associated modes evanesce. Inspection of the contour data shows that all eigenvalues lie in this shaded region if $k H<2.64$. This means that $k H=2.64$ is the minimum frequency at which a signal can propagate.


Figure 5.
The dashed curves in Fig. 5 marked $F_{\mathrm{pl}}=0$ are the locus of points at which the term within braces in both characteristic equations vanishes. This term is

$$
F_{\mathrm{pl}}=\frac{D}{\rho_{p} h H^{2} c^{2}} \frac{\left[(k H)^{2}-\eta^{2}\right]^{2}}{(k H)^{2}}-1
$$

This factor is the nondimensional version of $D \kappa^{4}-\rho_{p} \omega^{2}$ after substitution of $c^{2} \kappa^{2}=$ $\omega^{2}-c^{2} \mu^{2}$. In other words it is the dispersion equation for waves in an elastic plate that is not loaded by the fluid. The intersection of an $F_{\mathrm{pl}}$ curve with one of the eigenvalue contours corresponds to the coincidence frequency for a plate bounding a half-space. When $F_{\mathrm{pl}}=0$, the inertia and stiffness effects of the plate cancel. In the vicinity of the lower curve, the value of $F_{\mathrm{pl}}$ is relatively small, so the eigenvalues are relatively insensitive to this term. The upper curve is the contrary condition. The intersection of this $F_{\mathrm{pl}}$ curve with any eigenvalue curve is centered in the region where each curve is multivalued, which suggests that these values are very sensitive to the plate's properties. Note that the axial phase speed is $\omega / \kappa$, so the lower $F_{\mathrm{pl}}$ curve describes waves in the plate/wall that are supersonic.

A lesson that can be gleaned from this example is that systems in which there is a strong interaction between an acoustic medium and a structure might exhibit unexpected effects. We also have seen that the cause of these effects can be identified if the properties of the system's response are examined from an appropriate perspective. It is worth noting that the steps leading to the governing equations are typical of all systems in which elastic and acoustic subsystems strongly interact. Continuity of the normal displacement at the interface leads to a boundary condition for the fluid, and the pressure applied to the structure is accounted for as a dynamic load.

### 9.2.5 Orthogonality and Signal Generation

The primary motivation for identifying the modes of a waveguide is that they simplify an analysis of the signal that is generated by an excitation. This excitation is taken to be a specification of the axial particle velocity on the cross section at $x=0$. The velocity distribution is $V_{s}(y)$, so the boundary condition is

$$
\begin{equation*}
V_{x}=-\frac{1}{i \omega \rho_{0}} \frac{\partial P}{\partial x}=V_{s}(y) \text { at } x=0 \tag{9.2.42}
\end{equation*}
$$

The other condition is that waves are assumed to be absorbed or perfectly transmitted at a long range, so we are solely concerned with waves that propagate in the positive $x$ direction.

How to satisfy the source condition is the question addressed here. The situation is actually close to that which arose for a spherical source. There the signal propagated to an infinite radial distance, and we used a spherical harmonic series to represent the transverse variation. Here the propagation is in the axial direction. We know how the signal in a mode varies in the transverse $y$ direction, so it is logical to consider a modal series,

$$
\begin{equation*}
P=\sum_{n=0}^{\infty} P_{n} \Phi_{n}(y) e^{-i \kappa_{n} x} \tag{9.2.43}
\end{equation*}
$$

In this expression the $P_{n}$ coefficients are called the modal amplitudes of the mode functions. The sum is indicated to begin at $n=0$, but that is reserved for the plane wave mode in the case of a hard-walled waveguide. Also, if the waveguide has modes whose eigenvalues are purely imaginary, they must be inserted into this series.

The path leading to an expression for the modal amplitudes is like that by which the coefficients of a Fourier series are determined. Indeed, some individuals refer to a modal series, Eq. (9.2.43), as a generalized Fourier series. Both series are a sum of a basis function, modal or trigonometric, multiplied by an unknown coefficient. Each function represents a direction in a generalized linear space and the coefficient represents a vector component in that direction. The similarity extends further, in that the basis functions form an orthogonal set.

The procedure for deriving the orthogonality property is part of Sturm-Liouville analysis in a course on partial differential equations. We begin by observing that the transverse modes satisfy the mixed boundary conditions at $y=0$ and $y=H$ stated in Eq. (9.2.29),

$$
\begin{gather*}
\Phi_{n}=\frac{\zeta_{0}}{i k} \frac{d \Phi_{n}}{d y} \text { at } y=0 \\
\Phi_{n}=-\frac{\zeta_{H}}{i k} \frac{d \Phi_{n}}{d y} \text { at } y=H \tag{9.2.44}
\end{gather*}
$$

Within the domain of the fluid the mode functions satisfy a one-dimensional Helmholtz equation,

$$
\begin{equation*}
\frac{d^{2}}{d y^{2}} \Phi_{n}+\mu_{n}^{2} \Phi_{n}=0 \tag{9.2.45}
\end{equation*}
$$

The key step is selection of two arbitrary mode indices $n$ and $j$. The Helmholtz equation for each equation is multiplied by the other mode function. The difference of the two products is

$$
\begin{equation*}
\Phi_{j}\left(\frac{d^{2}}{d y^{2}} \Phi_{n}+\mu_{n}^{2} \Phi_{n}\right)-\Phi_{n}\left(\frac{d^{2}}{d y^{2}} \Phi_{j}+\mu_{j}^{2} \Phi_{j}\right)=0 \tag{9.2.46}
\end{equation*}
$$

The next operation entails integrating this expression over $0<y<H$,

$$
\begin{equation*}
\int_{0}^{H}\left(\Phi_{j} \frac{d^{2}}{d y^{2}} \Phi_{n}-\Phi_{n} \frac{d^{2}}{d y^{2}} \Phi_{j}\right) d y+\left(\mu_{n}^{2}-\mu_{j}^{2}\right) \int_{0}^{H} \Phi_{j} \Phi_{n} d y=0 \tag{9.2.47}
\end{equation*}
$$

Integration by parts is applied to the integral containing derivatives. Let us examine one of those terms,

$$
\begin{align*}
\int_{0}^{H} \Phi_{j} \frac{d^{2}}{d y^{2}} \Phi_{n} d y \equiv & -\int_{0}^{H} \frac{d}{d y} \Phi_{j} \frac{d}{d y} \Phi_{n} d y+\left.\left(\Phi_{j} \frac{d}{d y} \Phi_{n}\right)\right|_{y=H}  \tag{9.2.48}\\
& -\left.\left(\Phi_{j} \frac{d}{d y} \Phi_{n}\right)\right|_{y=0}
\end{align*}
$$

If a boundary is rigid, then the normal derivative is zero, whereas the function is zero if a boundary is pressure-release. In either case the terms that are evaluated at the boundaries vanish. Otherwise the impedance boundary conditions in Eq. (9.2.44) relate the $y$ derivative to the function at each wall. Their substitution into the preceding gives

$$
\begin{align*}
\int_{0}^{H} \Phi_{j} \frac{d^{2}}{d y^{2}} \Phi_{n} d y= & -\int_{0}^{H} \frac{d}{d y} \Phi_{j} \frac{d}{d y} \Phi_{n} d y+\left.\left(\frac{i k}{\zeta_{H}} \Phi_{j} \Phi_{n}\right)\right|_{y=H}  \tag{9.2.49}\\
& -\left.\left(\frac{i k}{\zeta_{0}} \Phi_{j} \Phi_{n}\right)\right|_{y=0}
\end{align*}
$$

When the same set of operations is performed on the other derivative term in Eq. (9.2.47), the result has the form of this relation with $n$ and $j$ swapped. However, such an interchange does not alter the result, which means that all terms cancel. Thus, regardless of the wall impedances, the integral in Eq. (9.2.47) that contains derivatives vanishes. What remains is the orthogonality property,

$$
\begin{equation*}
\left(\mu_{n}^{2}-\mu_{j}^{2}\right) \int_{0}^{H} \Phi_{j} \Phi_{n} d y=0 \tag{9.2.50}
\end{equation*}
$$

If $j$ and $n$ are different, then the integral must vanish. If the indices are equal, then the integral will not evaluate to zero. We can use this fact to set the arbitrary coefficient $B_{n}$ in Eq. (9.2.33). In particular we will select it such that the integral equals $H$, which will make the transverse mode functions dimensionless. Let $\tilde{\Phi}_{n}$ denote a transverse mode function when $B_{n}=1$, so that

$$
\begin{gather*}
\Phi_{n}=B_{n} \tilde{\Phi}_{n}  \tag{9.2.51}\\
B_{n}=\left[\frac{1}{H} \int_{0}^{H}\left(\tilde{\Phi}_{n}\right)^{2} d y\right]^{-1 / 2}
\end{gather*}
$$

The integral property for any pair of $n$ and $j$ values is captured by using a Kronecker delta, $\delta_{n, j}=0$ if $j \neq n$ and $\delta_{n, j}=1$ if $j=n$. Thus,

$$
\begin{equation*}
\int_{0}^{H} \Phi_{j} \Phi_{n} d y=H \delta_{n, j} \tag{9.2.52}
\end{equation*}
$$

From a linear algebra perspective the integral is the scalar product of $\Phi_{n}$ and $\Phi_{j}$ in the functional space of a single Cartesian coordinate. Correspondingly, we say that the mode functions form an orthonormal set. Without the above definition of $B_{n}$, the mode functions would be the analog of orthogonal vectors that have arbitrary magnitude. With it, the mode functions are analogous to a set of orthogonal unit vectors. A rule that scales the arbitrary coefficient of a mode according to a specified rule is said to be normalization. To some extent it would be less confusing to say that the rule "standardizes" a mode because "normal" is synonymous with "orthogonal", which is a property the modes always have. Nevertheless, the term is generally accepted, so we shall use it.

To determine the modal coefficients. we express the given boundary excitation as a series whose basis functions are the transverse modes $\Phi_{n}(y)$, specifically

$$
\begin{equation*}
\sum_{n=0}^{\infty} V_{n} \Phi_{n}(y)=V_{s}(y) \tag{9.2.53}
\end{equation*}
$$

An explicit description of the velocity coefficients is obtained from the modal orthogonality property. Thus, the preceding is multiplied by a specific function $\Phi_{j}(y)$, and
both sides are integrated over the cross section. This operation filters the $j$ th term from the sum, with the result that

$$
\begin{equation*}
V_{j}=\frac{1}{H} \int_{0}^{H} V_{s} \Phi_{j} d y \tag{9.2.54}
\end{equation*}
$$

At the same time, Euler's equation states that the axial particle velocity at $x=0$ corresponding to the modal pressure series is

$$
\begin{equation*}
V_{s}=\left.\frac{i}{\omega \rho_{0}} \frac{d P}{d x}\right|_{x=0}=\frac{1}{\omega \rho_{0}} \sum_{n=0}^{\infty} \kappa_{n} P_{n} \Phi_{n}(y) \tag{9.2.55}
\end{equation*}
$$

This description of the axial velocity must be the same as Eq. (9.2.53). The alternatives are modal series with the same orthogonal set of basis functions. Equality requires that the coefficients match, which leads to

$$
\begin{equation*}
P_{j}=\rho_{0} c \frac{k}{\kappa_{j}} V_{j} \tag{9.2.56}
\end{equation*}
$$

The corresponding expression for the pressure field is

$$
\begin{equation*}
P(x, y)=\rho_{0} c \sum_{j=0}^{\infty} \frac{k}{\kappa_{j}} V_{j} \Phi_{j}(y) e^{-i \kappa_{j} x} \tag{9.2.57}
\end{equation*}
$$

It is implicit to this expression that evaluation of $\kappa_{n}$ is based on a branch cut for the square root that places $\kappa_{n}$ in the fourth quadrant of the complex plane.

Equation (9.2.56) tells us that the source might generate an infinite number of waveguide modes. The degree to which the source velocity matches a mode function dictates the magnitude of the velocity coefficient. The other factor affecting the pressure coefficients is $\kappa_{n}$, whose minimum magnitude at a specific frequency occurs in the vicinity of $n_{\text {cutoff }}$. For $n \gg n_{\text {cutoff }}, \kappa_{n}$ will be much larger than $k$, which is one reason that the pressure series converges. This property leads to convergence of the modal series at any $x$. Convergence improves as $x$ increases because the evanescent modes decay increasingly.

The case of a point source situated somewhere on the cross section at $x=0$ leads to an explicit description. Recall that the volume velocity is the limit as the source's surface area shrinks to zero and the normal velocity of the source increases commensurately. Accordingly, the source is described by a Dirac delta function. For a two-dimensional waveguide with the source situated at $y=y_{0}$, the volume velocity is per unit width in the invariant direction. In other words, we are considering a line source. Hence, we have

$$
V_{s}=\lim _{\varepsilon \rightarrow 0}\left\{\begin{array}{l}
(\hat{Q} / \varepsilon) \text { if } y_{0}<y<y_{0}+\varepsilon  \tag{9.2.58}\\
0 \text { otherwise }
\end{array}\right\}=\hat{Q} \delta\left(y-y_{0}\right)
$$

The velocity coefficients obtained by application of Eq. (9.2.54) are $V_{j}=\hat{Q} \Phi_{j}$ $\left(y_{0}\right) / H$. The corresponding pressure field is

$$
\begin{equation*}
P_{\text {source }}=\rho_{0} c \frac{\hat{Q}}{H} \sum_{j=0}^{\infty} \frac{k}{\kappa_{j}} \Phi_{j}\left(y_{0}\right) \Phi_{j}(y) e^{-i \kappa_{j} x} \tag{9.2.59}
\end{equation*}
$$

It is evident that if the source is situated at a location where mode $\# j$ is largest, that mode will be excited to the greatest extent possible. Conversely, if the source is situated at a node of the $j$ th mode, that mode will not be excited. Furthermore, if the pressure is observed at a location where the $j$ th mode has a node, there will no evidence of that mode at that location, even though it might be an important mode for the pressure at other points.

Orthogonality of transverse mode functions plays an important role in an analysis of axial power flow. To evaluate the axial component of intensity we require an expression for $V_{x}$, which we obtain by applying Euler's equation to the modal series for pressure. This operation yields

$$
\begin{equation*}
V_{x}=\sum_{j=0}^{\infty} V_{j} \Phi_{j}(y) e^{-i \kappa_{j} x} \tag{9.2.60}
\end{equation*}
$$

The time-averaged power flow past any cross section is the average intensity at fixed $x$ integrated over $y$,

$$
\begin{align*}
\mathcal{P}_{\mathrm{av}}(x) & =\frac{1}{2} \operatorname{Re}\left\{\int_{0}^{H}\left[\rho_{0} c \sum_{j=0}^{\infty} \frac{k}{\kappa_{j}} V_{j} \Phi_{j}(y) e^{-i \kappa_{j} x}\right]\left[\sum_{n=0}^{\infty} V_{n}^{*} \Phi_{n}(y) e^{i \kappa_{n}^{*} x}\right] d y\right\} \\
& =\frac{\rho_{0} c}{2} \operatorname{Re}\left\{\sum_{j=0}^{\infty} \sum_{n=0}^{\infty} \frac{k}{\kappa_{j}} V_{j} V_{n}^{*} e^{-i\left(\kappa_{j}-\kappa_{n}^{*}\right) x} \int_{0}^{H} \Phi_{j}(y) \Phi_{n}(y) d y\right\} \tag{9.2.61}
\end{align*}
$$

Because the transverse mode functions form an orthogonal set that is normalized to $H$, the terms in the double sum for which $n \neq j$ integrate to zero. This reduces the expression to

$$
\begin{equation*}
\mathcal{P}_{\text {av }}(x)=\frac{\rho_{0} c H}{2} \operatorname{Re}\left[\sum_{j=0}^{\infty} \frac{k}{\kappa_{j}}\left|V_{j}\right|^{2} e^{2 \operatorname{Im}\left(\kappa_{j}\right) x}\right] \tag{9.2.62}
\end{equation*}
$$

The first notable aspect of this expression is that the modes do not couple, so the power flow is the sum of the contribution of each mode. If both walls are purely reactive, then the $\kappa_{j}$ values below cutoff ( $j \leq N_{\text {cutoff }}$ ) will be real. The exponential decay factor in that case is unity, which means that these modal contributions to the power flow are independent of $x$. The axial wavenumber is negative imaginary for
$j>N_{\text {cutoff. }}$. The exponential factor in the summation is real for these modes, but the presence of $\kappa_{j}$ in the denominator means that each of these terms is imaginary. The consequence is that evanescent modes do not contribute to the power flow. (The fundamental reason for this property is that the axial velocity in an evanescent mode is $90^{\circ}$ out of phase from the pressure.) This means that if there is no dissipation the power input to a waveguide at one end is transported downstream by the propagating modes without loss. The situation is different if the impedance of either wall has a resistive part. In that case all axial wavenumbers are complex. We established previously that the $\kappa_{j}$ values must lie in the fourth quadrant of the complex plane, which means $\operatorname{Im}\left(\kappa_{j}\right)<0$. Thus the power input to each mode at one end of the waveguide is progressively dissipated through the walls(s) as the mode propagates downstream. The decay rate is twice as fast as the decay of the pressure amplitude. If the waveguide is sufficiently long, the pressure and therefore the power, that reaches the far end will be negligible.

EXAMPLE 9.6 The sketch shows a rectangular tank of water that is open to the atmosphere. On the left wall a ribbon transducer is mounted at mid-depth. The transducer executes a uniform translational oscillation in the direction normal to the wall. Its width is sufficiently narrow that it may be approximated as a line whose a volume velocity per unit length is $\operatorname{Re}\left(\hat{Q}_{0} e^{i \omega t}\right)$. The bottom and side walls are composed of a thick glass sheet that may be considered to be rigid. The far wall consists of a mosaic of piezoceramic tiles that are actively controlled to achieve a nonreflecting boundary condition. The tank dimensions are $H=600 \mathrm{~mm}$ and $L=3 \mathrm{~m}$, the frequency is 20 kHz , and $\hat{Q}_{0}=8\left(10^{-4}\right)$ $\mathrm{m}^{2} / s$. The wall containing the ribbon transducer is the plane $x=0$, and $y$ is measured from the bottom. Determine the dependence of $|p|$ on the axial distance $x$ at mid-depth, $y=H / 2$, and also determine the depth dependence of $|p|$ at $x=0, L / 2$, and $L$. At each location, identify the portion that is attributable to the modes that are below the cutoff wavenumber. In addition, determine the spatially averaged impedance required of the piezoceramic tiles.


Figure 1.

## Significance

The best way to realize the effect of the cutoff phenomenon and the role of the evanescent spectrum, is to carry out a computation. The results will illustrate the general nature of multimodal signals in a waveguide.

## Solution

It might not be obvious that this system constitutes a two-dimensional waveguide because the tank is not infinitely wide in any direction parallel to the plane of the excitation. However, the vertical side walls are rigid, which means that $\partial P / \partial z$ must be zero along them. A pressure field that does not depend on $z$ will satisfy this condition. The ribbon transducer executes the same vibration at all $z$, so a pressure field fitting the description $\operatorname{Re}\left(P(x, y) e^{i \omega t}\right)$ is fully consistent. Thus, the pressure is described by the modal series in Eq. (9.2.43). The bottom of the tank is rigid, and the top is the free surface, which is well represented as pressure-release. The modes for such a configuration are described by Eq. (9.2.35),

$$
\begin{equation*}
\Phi_{n}=B_{n} \cos \left(\mu_{y} y\right), \quad \mu_{n}=\left(n-\frac{1}{2}\right) \frac{\pi}{H}, \quad n=1,2, \ldots \tag{1}
\end{equation*}
$$

The $B_{n}$ coefficients are normalized according to Eq. (9.2.51), which gives

$$
\begin{equation*}
B_{n}=H^{1 / 2}\left[\int_{0}^{H} \cos \left(\frac{(2 n-1)}{2} \frac{\pi y}{H}\right)^{2} d y\right]^{-1 / 2}=2^{1 / 2} \tag{2}
\end{equation*}
$$

The excitation is a line source having a specified velocity. Aside from the line at $y=H / 2$, the wall is rigid, so the velocity boundary condition at $x=0$ is

$$
\begin{equation*}
\left.V_{x}\right|_{x=0}=\hat{Q}_{0} \delta\left(y-\frac{H}{2}\right), \quad 0<y<H \tag{3}
\end{equation*}
$$

According to Eq. (9.2.54), the coefficients of a modal series for this velocity distribution are

$$
\begin{equation*}
V_{n}=\frac{1}{H} \int_{0}^{H} \hat{Q}_{0} \delta(y-H / 2) \Phi_{n} d y=\frac{\hat{Q}_{0}}{H} \Phi_{n}(y=H / 2)=\frac{\hat{Q}_{0}}{H} 2^{1 / 2} \cos \left(\frac{(2 n-1) \pi}{4}\right) \tag{4}
\end{equation*}
$$

The resulting modal series obtained from Eq. (9.2.57) is

$$
\begin{equation*}
P(x, y)=\rho_{0} c \sum_{n=0}^{\infty} \frac{k}{\kappa_{n}} V_{n} \Phi_{n}(y) e^{-i \kappa_{n} x} \tag{5}
\end{equation*}
$$

The last consideration before we may evaluate the pressure is selection of the number of terms $N$ to include in the series. The transverse wavenumbers are given
by Eq. (1). The cutoff condition is marked by $\mu_{n}=k=84.91$ based on $c=1480 \mathrm{~m} / \mathrm{s}$. This leads to

$$
N_{\mathrm{cut}}=\text { floor }\left(\frac{k H}{\pi}+\frac{1}{2}\right)=16
$$

As was noted, convergence of the modal series, Eq. (9.2.43), can result from reduction of $\left|P_{n}\right|$ to a negligible value above some $n$. This possibility constitutes the convergence criterion for small $x$, because $\left|\exp \left(-i \kappa_{n} x\right)\right|$ will be $O(1)$ regardless of whether $\kappa_{n}$ is real or negative imaginary. An important aspect is the fact that $\kappa_{n}$ becomes very small near $n_{\text {cutoff }}$. Because $\kappa_{n}$ appears in the denominator of Eq. (3), achieving convergence in this manner requires that $N>N_{\text {cut }}$. The other terms in Eq. (3) do not grow with increasing $n$, so setting $N$ substantially larger than $n_{\text {cutoff }}$ should be adequate. We shall use $N=3 N_{\text {cut }}$, which leads to $\left|P_{n}\right|<0.07 \mathrm{max}\left|P_{n}\right|$ for $n>N$.

The other attribute that leads to convergence of a modal series is the decay of the evanescent modes with increasing $x$. This decay increases with increasing mode number. Thus, fewer and fewer terms are required as $x$ increases. We could use this attribute to reduce $N$ at each $x$ based on the value of $\exp \left(-i \mu_{n} x\right)$ for $n>n_{\text {cutoff }}$. We shall not do so to avoid complicating the computational program. Furthermore, the processing time required for computations at fixed $N$ is not prohibitive. The only remaining consideration is a caution against a common error. The erroneous procedure evaluates the magnitude of each term in the modal series. The correct procedure evaluates the sum first.

Before we examine data, it is useful to recall the earlier matrix based algorithm for evaluation of a series. Our interest is in the values of $P$ at a set of $x$ and $y$ locations. In Eq. (5) the mode function depends on $y$, and the exponential function depends on $x$. Let these locations be $x_{j}, j=1, \ldots, J$ and $y_{m}, m=1, \ldots, M$. The summation over the mode number $n$ can be expressed as a matrix sum. To do so, let $[F]$ be a matrix of transverse mode functions and [ $E$ ] be a matrix of exponentials, such that $F_{n, m} \equiv \Phi_{n}\left(y_{m}\right)$ and $E_{j, n} \equiv \exp \left(-i \mu_{n} x_{j}\right)$. Then $P_{m, j} \equiv P\left(x_{j}, y_{m}\right)$ is the element of a rectangular array $[P]$ at all locations that is computed from the matrix version of Eq. (4), which is

$$
\begin{equation*}
[P]_{J \times M}=\rho_{0} c[E]_{J \times N}[K]_{N \times N}[F]_{N \times M} \tag{6}
\end{equation*}
$$

where [ $K$ ] is a diagonal matrix whose elements are $K_{n, n}=\left(k / \kappa_{n}\right) V_{n}$. A row of [ $P$ ] will be the pressure values at all $y_{m}$ for fixed $x_{j}$, whereas a column will be the values at all $x_{j}$ for fixed $y_{m}$. A bonus of this formulation is that changing the modal index to extend from $n=1$ to $n=N_{\text {cutoff }}$ is the only modification required to exclude the evanescent modes.

The first set of results we shall examine are the pressure amplitudes. According to Eq. (5), they are

$$
\hat{P}_{n}=\rho_{0} c \frac{k}{\kappa_{n}} V_{n}
$$

These values are described by Fig. 2. As was anticipated, the largest coefficients occur in the vicinity of $N_{\text {cut }}$.


Figure 2.
The axial dependence of the pressure may be computed according to Eq. (6) with $x_{j}$ being a set of axial locations and $y_{1}=H / 2$. Figure 3 shows the result. Many individuals anticipate a smooth curve, whereas the actual result shows rounded maxima separated by narrow near-null regions. Capturing these features requires a fine scan; the data for the Fig. 3, as well as for the computations along a cross section used a point spacing set at $\Delta=(2 \pi / k) / 25$. The cause of this behavior is interference, as it is for a sound beam. The interference occurs here because the multiple modes have different wavelengths. At some locations many are nearly in-phase creating maxima, whereas at other locations the modal contributions nearly annihilate each other. Another noteworthy feature is that the pressure very close to the transducer is 10 dB higher than it is at any downstream maximum.


Figure 3 shows that only very close to $x=0$ is there significant difference between the pressure obtained from the full series and the series that consists of only propagating modes. This observation is confirmed by the transverse profiles in Fig. 4. To obtain the data for Eq. (6), $x_{j}$ was set to the axial position and $y_{m}$ was a sequence of values spaced at $\Delta$. Interference is evident at each cross section. The cross section at the boundary, $x=0$, shows a sharp pressure peak in the vicinity of $y=0.3 \mathrm{~m}$, which is the location of the transducer. With increasing distance from the excitation, that tendency is overcome by the interference effect. Note that the profiles at $x=1.5 \mathrm{~m}$ and $x=3 \mathrm{~m}$ are similar in magnitude, as well as spatial dependence, but they are not the same. Further only at $x=0$ is the contribution of the evanescent modes significant.


Figure 4.
The last quantity to evaluate is the average impedance at $x=L$. This quantity is merely the average complex pressure amplitude divided by the average particle velocity. There is no variation in the $z$ direction, so averages per unit width are obtained by integrating over $y$. We use the modal pressure series in Eq. (5) to describe $P$ as a function of $y$, so that

$$
\begin{align*}
\left.P_{\mathrm{av}}\right|_{x=L} & =\frac{1}{H} \int_{0}^{H} P(L, y) d y=\frac{\rho_{0} c}{H} \sum_{n=1}^{\infty} \frac{k}{\kappa_{n}} V_{n} \int_{0}^{H} 2^{1 / 2} \cos \left(\frac{(2 n-1) \pi y}{2 H}\right) e^{-i \kappa_{n} L} d y \\
& =\rho_{0} c \sum_{n=0}^{\infty} \frac{k}{\kappa_{n}}\left(\frac{2^{3 / 2} V_{n}}{(2 n-1) \pi}\right)(-1)^{n-1} e^{-i \kappa_{n} L} \tag{7}
\end{align*}
$$

The average axial velocity is obtained by integrating Eq. (9.2.60) over $y$,

$$
\begin{align*}
\left.\left(V_{x}\right)_{\mathrm{av}}\right|_{x=L} & =\frac{1}{H} \int_{0}^{H} V_{x}(L, y) d y=\frac{1}{H} \sum_{n=1}^{\infty} V_{n} e^{-i \kappa_{n} L} \int_{0}^{H} 2^{1 / 2} \cos \left(\frac{(2 n-1) \pi y}{2 H}\right) d y \\
& =\sum_{n=1}^{\infty}\left(\frac{2^{3 / 2} V_{n}}{(2 n-1) \pi}\right)(-1)^{n-1} e^{-i \kappa_{n} L} \tag{8}
\end{align*}
$$

Despite the similarities of the sums in Eqs. (7) and (8), each sum must be evaluated individually. The average specific impedance that results from these expressions is

$$
\zeta_{\mathrm{av}}=\frac{Z_{\mathrm{av}}}{\rho_{0} c}=\frac{\left.P_{\mathrm{av}}\right|_{x=L}}{\left.\rho_{0} c\left(V_{x}\right)_{\mathrm{av}}\right|_{x=L}}=1.086-0.0043136 i
$$

The average impedance is very close to $\zeta=1$, which is the value for which the reflection coefficient of a plane wave is zero.

Before we leave this system, it is interesting to contemplate an alternative solution based on the method of images. Example 6.8 developed such a solution for the
case of a point source in a hard-walled waveguide. The ribbon transducer in the present system effectively is a line source, so all images also are line sources. Images above the free surface, $y=H$, invert images below that surface, while images below the bottom are in-phase replicas of the images above that surface. This creates an infinite sequence of line sources lying in the $x=0$ cross section centered at $y=H / 2$ and separated by distance $H$. The pressure amplitude radiated by a line source decays as $1 / R^{1 / 2}$ when the distance $R$ transverse to that line is large. If we were to impose a comparable convergence criterion to that employed for the modal series, we would cease a summation of all images when the $n$th line source above or below has an amplitude that is $1 \%$ of the largest contribution. For a field on the axis, the closest image is the actual transducer, which is at $y=H / 2, x=0$. The corresponding source-image distance is $R_{0}=x$. The distance to the same field point from the $n$th image is $R_{n}=\left[x^{2}+(n H)^{2}\right]^{1 / 2}$. The aforementioned convergence condition suggest that we may omit from the image summation any image for which $1 / R_{n}^{1 / 2}<0.01 / R_{0}^{1 / 2}$. This simplifies to $\left[x^{2}+(n H)^{2}\right]^{1 / 2}>10^{4} x$, which is well approximated as $n>10^{4} x / H$. In other words, the number of images required to obtain a suitably convergent representation of the on-axis pressure must increase drastically in proportion to the distance. The signal arriving at a field point from each image corresponds to a ray. In general, the number of rays required to represent a signal increases greatly with increasing range. In contrast, a fixed, relatively small number of modes usually suffices.

### 9.3 Three-Dimensional Waveguides

The two-dimensional model of a waveguide encapsulates the conceptual and phenomenological features of other waveguides whose cross section is invariant. This will be evident in the following general formulation and its application to cases of rectangular and circular cross sections.

### 9.3.1 General Analytical Procedure

As we did previously, we designate the direction in which signals propagate as $x$. The cross section is $\mathcal{A}$, whose shape and size are constant. The other restriction imposed on this general exploration is that the side wall, which forms the perimeter $\mathcal{C}$, is locally reacting. The specific impedance $\zeta$ may vary along the perimeter, but it is independent of $x$ (e.g., a rectangular cross section could have walls different types of sides). Let $\xi_{2}$ and $\xi_{3}$ denote the transverse coordinates. These would be Cartesian coordinates if the shape is rectangular, or polar coordinates if the shape were circular, but their definition is unimportant at this juncture. The transverse mode functions therefore are $\Phi_{n}\left(\xi_{2}, \xi_{3}\right)$, and a waveguide mode is obtained by multiplying these
functions by a complex exponential representing the propagation. In other words, we anticipate that the waveguide modes will be described as

$$
\begin{equation*}
P=\Phi_{n}\left(\xi_{2}, \xi_{3}\right) e^{-i \kappa_{n} x} \tag{9.3.1}
\end{equation*}
$$

As always, the wavenumber must lie in the fourth quadrant of the complex plane in order to satisfy the Sommerfeld radiation condition. The transverse mode functions $\Phi_{n}$ are solutions of a Helmholtz equation whose general form is

$$
\begin{equation*}
\tilde{\nabla}^{2}\left(\Phi_{n}\right)+\mu_{n}^{2} \Phi_{n}=0, \quad \mu_{n}^{2}=k^{2}-\kappa_{n}^{2} \tag{9.3.2}
\end{equation*}
$$

where $\tilde{\nabla}^{2}$ denotes the portion of the Laplacian that contains derivatives with respect to $\xi_{2}$ and $\xi_{3}$.

As was the case for a two-dimensional waveguide, values of $\mu_{n}$ will be found from the condition that there be a nontrivial solution consistent with homogeneous boundary conditions at the walls. Application of Euler's equation to describe the particle velocity in the local impedance model, with $\bar{n}$ defined to be into the waveguide, leads to

$$
\begin{equation*}
P=\rho_{0} c \zeta(-\bar{n} \cdot \bar{V})=\frac{\zeta}{i k} \bar{n} \cdot \tilde{\nabla} P \text { if } \bar{x} \in \mathcal{C} \tag{9.3.3}
\end{equation*}
$$

where $\mathcal{C}$ is the perimeter of the cross section $\mathcal{A}$. Substitution of the representation of $P$ as a waveguide mode leads to a boundary condition for the transverse mode functions,

$$
\begin{equation*}
\Phi_{n}=\frac{\zeta}{i k} \bar{n} \cdot \tilde{\nabla} \Phi_{n} \tag{9.3.4}
\end{equation*}
$$

In addition to similarities of the equations that must be solved to determine the transverse mode functions, the general problem is like the two-dimensional one because its transverse mode functions are orthogonal. The general orthogonality condition is obtained similarly to the derivation of Eq. (9.2.52). Thus, we consider two eigensolutions and form

$$
\begin{align*}
& \Phi_{j}\left[\tilde{\nabla}^{2}\left(\Phi_{n}\right)+\mu_{n}^{2} \Phi_{n}\right]-\Phi_{n}\left[\tilde{\nabla}^{2}\left(\Phi_{j}\right)+\mu_{j}^{2} \Phi_{j}\right]  \tag{9.3.5}\\
& \quad \equiv \tilde{\nabla} \cdot\left(\Phi_{j} \tilde{\nabla} \Phi_{n}\right)-\tilde{\nabla} \cdot\left(\Phi_{n} \tilde{\nabla} \Phi_{j}\right)+\left(\mu_{n}^{2}-\mu_{j}^{2}\right) \Phi_{j} \Phi_{n}=0
\end{align*}
$$

This relation is integrated over a cross section. Application of the divergence theorem, with $\bar{n}$ defined as the normal on the perimeter $\mathcal{C}$ pointing into $\mathcal{A}$, leads to

$$
\begin{equation*}
\int_{C} \bar{n} \cdot\left(\Phi_{n} \tilde{\nabla} \Phi_{j}-\Phi_{j} \tilde{\nabla} \Phi_{n}\right) d \mathcal{C}+\left(\mu_{n}^{2}-\mu_{j}^{2}\right) \iint_{\mathcal{A}} \Phi_{j} \Phi_{n} h_{2}\left(\xi_{2}\right) h_{3}\left(\xi_{3}\right) d \mathcal{A}=0 \tag{9.3.6}
\end{equation*}
$$

The terms evaluated on $\mathcal{C}$ will vanish for the $\operatorname{rigid}\left(\bar{n} \cdot \tilde{\nabla} \Phi_{n}=0\right)$ and pressurerelease cases. If the wall is locally reacting, then Eq. (9.3.4) gives

$$
\begin{equation*}
\Phi_{n}\left(\bar{n} \cdot \tilde{\nabla} \Phi_{j}\right)=\Phi_{n}\left(\frac{i k}{\zeta} \Phi_{j}\right)=\left(\frac{i k}{\zeta} \Phi_{n}\right) \Phi_{j}=\left(\bar{n} \cdot \tilde{\nabla} \Phi_{n}\right) \Phi_{j} \tag{9.3.7}
\end{equation*}
$$

Thus the integral over $\mathcal{C}$ reduces to zero for any local impedance, even if $\zeta$ varies along $\mathcal{C}$.

What remains in Eq. (9.3.6) is the integral over $\mathcal{A}$. If the two mode functions correspond to different eigenvalues, $\mu_{j} \neq \mu_{n}$, then the integral must vanish. In the case where $\mu_{j}=\mu_{n}$, the integral may have any value. We shall use this arbitrariness to normalize the mode functions. Specifically, the $\Phi_{n}$ are scaled such that the integral equals the cross-sectional area. Thus, we will have

$$
\begin{equation*}
\iint_{\mathcal{A}} \Phi_{j} \Phi_{n} d \mathcal{A}=\mathcal{A} \delta_{j, n} \tag{9.3.8}
\end{equation*}
$$

Fulfillment of this condition when $j=n$ sets the $B_{n}$ coefficient of $\Phi_{n}$, which is arbitrary in regard to satisfying the eigenvalue problem.

Given the similarity of the general relations governing three-dimensional transverse mode functions and those for a two-dimensional waveguide, it is consistent that a series of these functions is used to represent the pressure in the general case. The coefficients of the modes are again designated as $P_{n}$, so that

$$
\begin{equation*}
P=\sum_{n=1}^{\infty} P_{n} \Phi_{n}\left(\xi_{2}, \xi_{3}\right) e^{-i \kappa_{n} x} \tag{9.3.9}
\end{equation*}
$$

The axial velocity distribution corresponding to this description of the pressure is

$$
\begin{equation*}
V_{x}=\sum_{n=1}^{\infty}\left(\frac{\kappa_{n}}{\kappa}\right) \frac{P_{n}}{\rho_{0} c} \Phi_{n}\left(\xi_{2}, \xi_{3}\right) e^{-i \kappa_{n} x} \tag{9.3.10}
\end{equation*}
$$

The procedure for determining the modal amplitudes is the same as it was for the two-dimensional waveguide. Typically, the excitation consists of a specified axial velocity $V_{s}\left(\xi_{2}, \xi_{3}\right)$ at $x=0$. This distribution is described by a modal series, specifically

$$
\begin{equation*}
V_{s}\left(\xi_{2}, \xi_{3}\right)=\sum_{n=1}^{\infty} V_{n} \Phi_{n}\left(\xi_{2}, \xi_{3}\right) \tag{9.3.11}
\end{equation*}
$$

To obtain an expression for the velocity coefficients the preceding is multiplied by an arbitrarily selected $\Phi_{j}$ and integrated over $\mathcal{A}$. Orthogonality, Eq. (9.3.8), filters the $j$ th term out of the summation, which leads to

$$
\begin{equation*}
V_{j}=\frac{1}{\mathcal{A}} \iint_{\mathcal{A}} \Phi_{j}\left(\xi_{2}, \xi_{3}\right) V_{s}\left(\xi_{2}, \xi_{3}\right) d \mathcal{A} \tag{9.3.12}
\end{equation*}
$$

Equation (9.3.10) evaluated at $x=0$ must be the same as the modal series of $V_{s}$. This requires that the coefficients of $\Phi_{n}$ match, which leads to

$$
\begin{equation*}
P_{n}=\rho_{0} c \frac{k}{\kappa_{n}} V_{n} \tag{9.3.13}
\end{equation*}
$$

There is no difference between this expression and Eq. (9.2.56) for the twodimensional system.

These expressions lead to a general description of the power flow. The timeaveraged power flowing across any cross section is the integral over that surface of the time-averaged intensity,

$$
\begin{align*}
\mathcal{P}(x)=\iint_{\mathcal{A}} \frac{1}{2} \operatorname{Re}\left(P V_{x}^{*}\right) d \mathcal{A}= & \frac{1}{2} \operatorname{Re} \iint_{\mathcal{A}}\left[\sum_{n=1}^{\infty} P_{n} \Phi_{n}\left(\xi_{2}, \xi_{3}\right) e^{-i \kappa_{n} x}\right] \\
& \times\left[\sum_{j=1}^{\infty}\left(\frac{\kappa_{j}^{*}}{\kappa}\right) \frac{P_{j}^{*}}{\rho_{0} c} \Phi_{j}\left(\xi_{2}, \xi_{3}\right) e^{+i \kappa_{j}^{*} x}\right] d \mathcal{A} \tag{9.3.14}
\end{align*}
$$

Different indices are used for the summations to assure that all products are included. The only quantities that depends on position are the $\Phi_{n}$ functions. This allows us to sum the integrals of the products of these functions. The result is that

$$
\begin{equation*}
\mathcal{P}(x)=\frac{1}{2} \operatorname{Re} \sum_{n=1}^{\infty} \sum_{j=1}^{\infty}\left(\frac{\kappa_{j}^{*}}{\kappa}\right) \frac{P_{n} P_{j}^{*}}{\rho_{0} c} e^{-i\left(\kappa_{n}-\kappa_{j}^{*}\right) x} \iint_{\mathcal{A}} \Phi_{n}\left(\xi_{2}, \xi_{3}\right) \Phi_{j}\left(\xi_{2}, \xi_{3}\right) d \mathcal{A} \tag{9.3.15}
\end{equation*}
$$

The last step is to apply the orthogonality of the transverse mode functions, which eliminates all coupling between modes. Thus, we find the power flow to be a sum of modal contributions,

$$
\begin{equation*}
\mathcal{P}(x)=\frac{\mathcal{A}}{2 \rho_{0} c} \operatorname{Re} \sum_{n=1}^{\infty} \frac{\kappa_{n}^{*}}{k}\left|P_{n}\right|^{2} e^{+2 \operatorname{Im}\left(\kappa_{n}\right) x} \tag{9.3.16}
\end{equation*}
$$

This is a useful relation because it allows for identification of modes that transport the greatest amount of energy. In the case of a purely reactive wall, the axial wave numbers are real for modes that propagate, and imaginary for evanescent modes. The exponential factor for the latter modes is real, but $\operatorname{Re}\left(\kappa_{n}^{*}\right)=0$, so the evanescent modes do not contribute the power transfer. Furthermore, $\operatorname{Im}\left(\kappa_{n}\right)=0$ for
the propagating modes, so the contribution of these modes to the power transfer is independent of $x$. Thus, $\mathcal{P}(x)$ is independent of $x$ if the walls are purely reactive, regardless of the shape of the cross section. The modal contributions to the power flow depend on $x$ only if the wall has a nonzero resistance. In that case the eigenvalues and, therefore the axial wavenumbers, are complex. We selected the branch cut for the square root such that $\operatorname{Im}\left(\kappa_{n}\right)<0$ in order to satisfy the radiation condition. The consequence is that power flow is attenuated in the downstream direction. The energy that is dissipated is absorbed in the wall.

The relations we have derived for the pressure, particle velocity, and power flow are generally applicable. However. their implementation assumes that we have identified transverse mode functions that satisfy the Helmholtz equation, Eq. (9.3.2), subject to the wall condition in Eq. (9.3.4). The ease of that determination depends on the shape of the cross section. The following studies explore rectangular and circular cross sections, which are the most amenable to formal analysis. Even within that limitation, an analysis using standard methods is not always possible. The derivation allowed the wall impedance to vary along the perimeter. However, the analyses we shall perform require that the boundary condition be constant along any segment of a wall that corresponds to a constant coordinate. In other words, the impedance of each side of a rectangular cross section is required to be constant, and the impedance of a circular wall cannot depend on the circumferential angle. If these conditions are not met, or if the subject is a waveguide whose cross section has an irregular shape, numerical techniques will be required to determine the transverse mode functions. Nevertheless, they do exist for such systems, and they do constitute an orthogonal set.

### 9.3.2 Rectangular Waveguide

A waveguide whose cross section is rectangular is the least challenging to analyze because we can define a Cartesian coordinate system whose coordinate planes are parallel to the walls. The walls in Fig. 9.13 correspond to $y=0, y=H, z=0$, $z=W$, and the excitation is applied to the end at which $x=0$. Waves propagate in the positive $x$ direction, and it is assumed that the far end is terminated in a manner that suppresses reflections.

Fig. 9.13 Definition of the Cartesian coordinate system used to analyze propagation in a rectangular waveguide


The waveguide mode is

$$
\begin{equation*}
P(x, y, z)=\Phi_{n}(y, z) \exp \left(-i \kappa_{n} x\right) \tag{9.3.17}
\end{equation*}
$$

The differential equation governing a transverse mode function $\Phi(y, z)$ is found by substituting this form into the Helmholtz equation, which leads to

$$
\begin{equation*}
\left(\frac{\partial^{2}}{\partial y^{2}}+\frac{\partial^{2}}{\partial z^{2}}\right) \Phi+\left(k^{2}-\kappa^{2}\right) \Phi_{n}=0 \tag{9.3.18}
\end{equation*}
$$

Each wall may have a local impedance $\zeta_{j}$. The normal directions into the fluid are $\bar{e}_{2}$ for the wall at $y=0,-\bar{e}_{2}$ for the wall at $y=H, \bar{e}_{3}$ for the wall at $z=0$, and $-\bar{e}_{3}$ for the wall at $z=W$, so the specific forms of the general local impedance boundary condition, Eq. (9.3.4), are

$$
\begin{align*}
P & =-\rho_{0} c \zeta_{1} V_{y}=\frac{\zeta_{1}}{i k} \frac{\partial P}{\partial y} \text { at } y=0 \\
P & =\rho_{0} c \zeta_{2} V_{y}=-\frac{\zeta_{2}}{i k} \frac{\partial P}{\partial y} \text { at } y=H  \tag{9.3.19}\\
P & =-\rho_{0} c \zeta_{3} V_{z}=\frac{\zeta_{3}}{i k} \frac{\partial P}{\partial z} \text { at } z=0 \\
P & =\rho_{0} c \zeta_{4} V_{z}=-\frac{\zeta_{4}}{i k} \frac{\partial P}{\partial z} \text { at } z=W
\end{align*}
$$

We could identify $\Phi$ by separation of variables, but it is more direct to guess the form of the general solution $\Phi$ and then seek the condition that such a form must satisfy. The analysis of a two-dimensional waveguide began by expressing $\Phi$ as a sum of complex exponentials, so it is logical to try the same here. Because either $y$ or $z$ is constant on a boundary, while the other coordinate has a range of values, a solution that is capable of simultaneously solving the boundary conditions on all walls must be a product of functions of $y$ and $z$. Thus, the appropriate ansatz is

$$
\begin{gather*}
\Phi=\phi_{y}(y) \phi_{z}(z) \\
\phi_{y}=B_{1 y} e^{-i \mu y}+B_{2 y} e^{i \mu y}, \quad \phi_{z}=B_{1 z} e^{-i \nu z}+B_{2 z} e^{i \nu z} \tag{9.3.20}
\end{gather*}
$$

The trial solution will satisfy the Helmholtz equation if

$$
\begin{equation*}
\mu^{2}+\nu^{2}+\kappa^{2}=k^{2} \tag{9.3.21}
\end{equation*}
$$

We substitute the trial form for $\Phi$ in Eq. (9.3.20) into the boundary conditions. The $\phi_{z}$ function is a common factor in the conditions at constant $y$, and $\phi_{y}$ is a common factor in the conditions on the walls at constant $z$. Cancelation of these common factors leads to two pairs of boundary conditions, which may be written as

$$
\begin{array}{|l}
{\left[D\left(\mu H, k H, \zeta_{1}, \zeta_{2}\right)\right]\left[\begin{array}{ll}
B_{1 y} & B_{2 y}
\end{array}\right]^{\mathrm{T}}=\left[\begin{array}{ll}
0 & 0
\end{array}\right]^{\mathrm{T}}}  \tag{9.3.22}\\
{\left[D\left(\nu W, k W, \zeta_{3}, \zeta_{4}\right)\right]\left[\begin{array}{ll}
B_{1 z} & B_{2 z}
\end{array}\right]^{\mathrm{T}}=\left[\begin{array}{ll}
0 & 0
\end{array}\right]^{\mathrm{T}}} \\
\hline
\end{array}
$$

Let $\alpha$ and $\beta$ denote the first two variables in the coefficient matrices for each direction, and $\zeta_{a}$ and $\zeta_{b}$ be the corresponding specific impedances. Then these matrices are described by

$$
\left[D\left(\alpha, \beta, \zeta_{a}, \zeta_{b}\right)\right] \equiv\left[\begin{array}{cc}
\left(\zeta_{a} \alpha+\beta\right) & \left(-\zeta_{a} \alpha+\beta\right)  \tag{9.3.23}\\
\left(-\zeta_{b} \alpha+\beta\right) e^{-i \eta \ell} & \left(\zeta_{b} \alpha+\beta\right) e^{i \alpha}
\end{array}\right]
$$

In this way the analysis reduces to finding the transverse modes for two twodimensional waveguides. The wall conditions at constant $y$ set the $\mu_{n}$ eigenvalues, and the conditions at constant $z$ set the $\nu_{n}$ values. The characteristic equations are

$$
\begin{equation*}
\left|\left[D\left(\mu H, k H, \zeta_{1}, \zeta_{2}\right)\right]\right|=0, \quad\left|\left[D\left(\nu W, k W, \zeta_{3}, \zeta_{4}\right)\right]\right|=0 \tag{9.3.24}
\end{equation*}
$$

The representations in Eq. (9.3.20) account for values of $\mu$ and $\nu$ that have opposite sign, so the only roots for which $\operatorname{Re}(\mu) \geq 0$ and $\operatorname{Re}(\nu) \geq 0$ should be sought.

The transverse mode function is denoted as $\Phi_{n, j}$ and referred to as the ( $n, j$ ) mode. Only the ratios $B_{2 y} / B_{1 y}$ and $B_{2 z} / B_{1 z}$ can be obtained from Eq. (9.3.22) because the respective [ $D$ ] matrix is rank-deficient when $\mu H$ is an eigenvalue. We use the first row of $[D]$ to determine these ratios. Furthermore, when the $y$ and $z$ functions are multiplied, only one arbitrary coefficient, which we denote as $B_{n}$, remains. The result is that the transverse mode functions may be written as a product of functions of $y$ and $z$, which we denote respectively as $\phi_{y, n}$ and $\phi_{z, j}$. Thus, we have established that

$$
\begin{equation*}
\Phi_{n, j}(y, z)=B_{n, j} \phi_{y, n}(y) \phi_{z, j}(z) \tag{9.3.25}
\end{equation*}
$$

where the factors are

$$
\begin{align*}
\phi_{y, n} & =\frac{\left(\zeta_{1} \mu_{n} H-k H\right) e^{-i \mu_{n} y}+\left(\zeta_{1} \mu_{n} H+k H\right) e^{i \mu_{n} y}}{\left(\zeta_{1} \mu_{n} H-k H\right)}  \tag{9.3.26}\\
\phi_{z, j} & =\frac{\left(\zeta_{3} \nu_{j} W-k W\right) e^{-i \gamma_{j} y}+\left(\zeta_{3} \nu_{j} W+k W\right) e^{i \nu_{j} y}}{\left(\zeta_{3} \nu_{j} W-k W\right)}
\end{align*}
$$

After the eigenvalues for each direction have been determined, the set of axial wavenumbers are

$$
\begin{equation*}
\kappa_{n, j}= \pm\left(k^{2}-\mu_{n}^{2}-\nu_{j}^{2}\right)^{1 / 2} \tag{9.3.27}
\end{equation*}
$$

where the alternative sign should be selected such that $\operatorname{Re}\left(\kappa_{n, j}\right)>0$ and $\operatorname{Im}\left(\kappa_{n, j}\right) \leq 0$. If all walls are reactive, then the $\mu_{n}$ and $\nu_{j}$ values are either real or purely imaginary. In such cases cutoff occurs if $\mu_{n}^{2}+\nu_{j}^{2}>k^{2}$. A corollary is that there is no single cutoff. Rather, for a fixed mode number in one direction there is a
mode number for the other direction beyond which the signal evanesces. For example, if $\mu_{5}=0.8 k$, then modes are cutoff in the $z$ direction if $\nu_{j}>0.6 k$. It follows that if $\mu_{n}>k$, then all modes at that $n$ and higher are evanescent, regardless of the value of $\nu_{j}$. Similarly, if $\nu_{j}>k$, then all modes at that $j$ and higher are evanescent. The alternative situation occurs when one or more wall has a resistance, so that its impedance has a real part. In such circumstances all modes decay axially, so that the demarcation between propagating and evanescent modes does not exist. It is reasonable to consider modes to be cutoff if the scale over which they decay, which is $-1 / \operatorname{Im}\left(\kappa_{n, j}\right)$, is comparable to or smaller than an axial wavelength, $2 \pi / \operatorname{Re}\left(\kappa_{n . j}\right)$. Such a transition will be a noticeable feature of dispersion curves if the resistive part of the wall impedance is small, as it was in Example 9.4. On the other hand, it might be meaningless to characterize a mode in this manner if the resistance is large.

Like a two-dimensional waveguide, a waveguide mode in a rectangular waveguide may be viewed either as a nonuniform plane wave that travels in the $x$ direction with the transverse dependence described by $\Phi_{n, j}$, or as a superposition of a set of plane waves that travel obliquely to the $x$-axis. Merging Eqs. (9.3.17) and (9.3.20) for a specific $(n, j)$ pair shows that a waveguide mode is a sum of four simple plane waves described by

$$
\begin{gather*}
\Phi_{n, j} e^{-i \kappa_{n, j} x}=\sum_{m=1}^{4} C_{m} e^{-i \bar{k}_{m} \cdot \bar{x}}  \tag{9.3.28}\\
\bar{k}_{m}=\kappa_{n, j} \bar{e}_{x} \pm \mu_{n, j} \bar{e}_{y} \pm \nu_{n, j} \bar{e}_{z}
\end{gather*}
$$

where the $C_{m}$ are set by satisfying the boundary conditions at the walls. If $\kappa_{n, j}$ is real, the four plane waves that combine to form a waveguide mode undergo multiple reflections at the side walls. Their interference pattern on a cross section produces the transverse mode function. This is the same phenomenon as the one illustrated in Fig. 9.9, except that each $\bar{k}_{m}$ has nonzero direction angles in both the $x y$ - and $x z$-planes.

The general orthogonality property, Eq. (9.3.8), in the case of a rectangular cross section states that

$$
\begin{equation*}
\int_{0}^{H} \int_{0}^{W} \Phi_{n, j} \Phi_{m, \ell} d z d y=H W \delta_{n . m} \delta_{j . \ell} \tag{9.3.29}
\end{equation*}
$$

The case where $m=n$ and $\ell=j$ sets the coefficient $B_{n, j}$ in Eq. (9.3.25). Substitution of that expression into the orthogonality condition leads to

$$
\begin{equation*}
B_{n, j}=(H W)^{1 / 2}\left(\int_{0}^{H}\left(\phi_{y, n}\right)^{2} d y\right)^{-1 / 2}\left(\int_{0}^{W}\left(\phi_{z, j}\right)^{2} d z\right)^{-1 / 2} \tag{9.3.30}
\end{equation*}
$$

We have seen that the factors $\phi_{y, n}$ and $\phi_{z, j}$ in Eq. (9.3.25) are transverse mode functions for a two-dimensional waveguide. This can be a very useful attribute, because it might be quite simple to form the transverse mode functions for a threedimensional system. For example, consider the three-dimensional transverse mode functions of the tank in Example 9.6. In the $y$ direction, it is a rigid/pressure-release
configuration, whereas it is a hard-walled waveguide in the $z$ direction. Multiplying a mode function for each gives

$$
\begin{equation*}
\Phi_{n, j}=B_{n, j} \cos \left(\frac{(2 n-1) \pi y}{2 H}\right) \cos \left(\frac{j \pi z}{W}\right), n=1,2, \ldots, j=0,1, \ldots \tag{9.3.31}
\end{equation*}
$$

A system of particular interest is a hard-walled waveguide, that is, one whose four walls are rigid. The combination of cosine functions in each direction gives

$$
\begin{equation*}
\Phi_{n, j}(y, z)=B_{n, j} \cos \left(\frac{j \pi y}{H}\right) \cos \left(\frac{n \pi z}{W}\right), \quad j, n=0,1,2, \ldots \tag{9.3.32}
\end{equation*}
$$

The corresponding eigenvalues and axial wavenumber are

$$
\begin{equation*}
\mu_{n}=\frac{n \pi}{H}, \quad \nu_{j}=\frac{j \pi}{W}, \quad \kappa_{j, n}= \pm\left[k^{2}-\left(\frac{n \pi}{H}\right)^{2}-\left(\frac{j \pi}{W}\right)^{2}\right]^{1 / 2} \tag{9.3.33}
\end{equation*}
$$

The alternative sign in this expression should be selected such that $\operatorname{Re}\left(\kappa_{n}\right)>0$ and $\operatorname{Im}\left(\kappa_{n}\right) \leq 0$.

The case where $j=n=0$ is the plane wave mode, in which $\Phi_{0,0}$ is independent of $y$ and $z$. The hard-walled waveguide is the only one in which a plane wave mode can exist. This is so for the same reason as a two-dimensional waveguide: In the plane wave mode $\bar{n} \cdot \nabla P$ is identically zero at the walls, which conflicts with the proportionality of pressure and velocity imposed by a finite local impedance.

Examination of Eq. (9.3.33) leads to recognition of an unusual feature. Suppose that the cross section is square, $W=H$. Then any $(n, j)$ mode for which $n \neq j$ has the same axial wavenumber as the $(j, n)$ mode. The same situation can arise even if $W \neq H$. To explore this possibility, let $W=r H$, and let $\left(j_{1}, n_{1}\right)$ and $\left(j_{2}, n_{2}\right)$ be the numbers for a pair of modes that have the same $\kappa$ value. Repeated values of $\kappa_{j, n}$ occur if

$$
\begin{equation*}
\left(n_{2}\right)^{2}+\left(\frac{j_{2}}{r}\right)^{2}=\left(n_{1}\right)^{2}+\left(\frac{j_{1}}{r}\right)^{2} \tag{9.3.34}
\end{equation*}
$$

This condition is met if $n_{2}=j_{1} / r$ and $j_{2}=r n_{1}$. Because the mode indices are integers, repeated $k_{n, j}$ values can arise only if $r$ is a rational number, that is, if $W / H$ is a ratio of integers. Furthermore, unless this ratio consists of small numbers, the overlapping modes will be very high, and possibly evanescent. For example if $r=3 / 2$, then $\kappa_{3,2}=\kappa_{2,3}, \kappa_{6,4}=\kappa_{4,6}$, etc. In contrast of $r=20 / 21$, then $\kappa_{20,21}=\kappa_{21,20}$, etc. Similar situations might arise for other combinations of rigid and pressure-release boundary conditions. Equality of the axial wavenumbers for different modes does not present a complication for the analysis. However, it can cause confusion for an examination of measured data.

After we have evaluated the transverse mode functions, the response to an excitation may be analyzed. The modal summation in Eq. (9.3.9) now must extend over the index for each direction, that is,

$$
\begin{equation*}
P=\sum_{n=0}^{\infty} \sum_{j=0}^{\infty} P_{n, j} \Phi_{n, j} e^{-i \kappa_{n, j} x} \tag{9.3.35}
\end{equation*}
$$

As always, a zero subscript applies only if a plane wave mode exists, in which case $\kappa_{n, j}=k$. Similarly, if there is an eigenvalue that is purely imaginary, its contribution must be inserted. The velocity excitation $V_{s}(x, y)$ extends over the cross section at $x=0$. Its modal series also is a double sum,

$$
\begin{equation*}
V_{s}=\sum_{n=0}^{\infty} \sum_{j=0}^{\infty} V_{n, j} \Phi_{n, j} \tag{9.3.36}
\end{equation*}
$$

The modes form an orthogonal set according to Eq. (9.3.29). This leads to the specific form of Eq. (9.3.12),

$$
\begin{equation*}
V_{n, j}=\frac{B_{n, j}}{H W} \int_{0}^{W} \int_{0}^{H} V_{s} \phi_{y, n} \phi_{z, j} d y d z \tag{9.3.37}
\end{equation*}
$$

These coefficients are matched to the modal series for $V_{x}$ at $x=0$ obtained from Euler's equation, as was done to obtain Eq. (9.3.13), with the result that

$$
\begin{equation*}
P_{n, j}=\rho_{0} c \frac{k}{\kappa_{n, j}} V_{n, j} \tag{9.3.38}
\end{equation*}
$$

The considerations regarding convergence for two-dimensional systems are equally applicable here. The largest $y$ mode number $N_{\text {cut }}$ for a propagating mode is such that $\mu_{N_{\text {cut }}}<k$ and $\mu_{N_{\text {cut }}+1}>k$. Similarly, the $z$ mode number for cutoff is such that $\nu_{J_{\mathrm{cut}}}<k$ and $\nu_{J_{\mathrm{cut}}+1}>k$. In regions that are not close to the excitation at $x=0$, truncation of the modal series at $n=N_{\text {cut }}$ and $j=J_{\text {cut }}$ will be adequate. This is so because the modes that are omitted correspond to $\kappa_{n, j}=-i \operatorname{Im}\left(\kappa_{n},{ }_{j}\right)$, so that $\exp \left(-i \kappa_{n, j} x\right)$ is negligible. In regions very close to the excitation at $x=0$, the evanescent modes have not decayed much. Convergence of the pressure series requires inclusion of the evanescent modes. The values of $\left|\kappa_{n, j}\right|$ increase rapidly if $n>N_{\text {cut }}+1$ or $j>J_{\text {cut }}+1$. A truncation at $N=2 N_{\text {cut }}$ and $j=2 J_{\text {cut }}$ should be adequate, unless $x$ is smaller than a wavelength. For such cases, truncation at $3 N_{\text {cut }}$ and $3 J_{\text {cut }}$ will suffice.

EXAMPLE 9.7 Consider the tank in Example 9.6 in the situation where the ribbon transducer is mounted on the vertical center line at $x=0$, rather than the horizontal center line. All parameters are as stated there, and $W=0.5$ m . Graph the dependence of $|P|$ on the axial distance along the center line $y=H / 2, z=W / 2$. Also, graph the profile of $|P|$ for cross sections at $x=0$ and $x=L$ along the horizontal line $y=H / 2$ and the vertical line $z=W / 2$.

## Significance

The number of modes is the product $N J$ of the series lengths, which might be quite large. Thus, it is desirable that the computational algorithm be efficient. The need to use two indices to specify a mode complicates development of such an algorithm. Both issues are addressed.

## Solution

We shall use the same coordinate system as that in Example 9.6. Unlike the previous arrangement, the ribbon transducer's velocity distribution varies in the $z$ direction. The consequence is that more than one mode in that direction is excited. Furthermore, although the transducer velocity now is independent of $y$, there is no plane wave mode for this direction. Consequently, all modes in that direction participate, as they did in the previous example. The $y$ dependence of the transverse mode function is the same rigid/pressure-release combination as it was in the previous Example, while the $z$ dependence consists of the modes of a two-dimensional hard-walled waveguide. Thus, the transverse mode functions are

$$
\Phi_{n, j}=B_{n, j} \cos \left(\frac{(2 n-1) \pi y}{2 H}\right) \cos \left(\frac{j \pi z}{W}\right), \quad n=1,2, \ldots, \quad j=0,1,2, \ldots
$$

The corresponding eigenvalues are

$$
\begin{equation*}
\mu_{n}=\frac{(2 n-1) \pi}{2 H}, \quad \nu_{j}=\frac{j \pi}{W} \tag{1}
\end{equation*}
$$

The axial wavenumber is either a positive real value or a negative imaginary value,

$$
\kappa_{n, j}=\left\{\begin{array}{l}
\left(k^{2}-\mu_{n}^{2}-\nu_{j}^{2}\right)^{1 / 2} \text { if } \mu_{n}^{2}+\nu_{j}^{2}<k^{2}  \tag{2}\\
-i\left(\mu_{n}^{2}+\nu_{j}^{2}-k^{2}\right)^{1 / 2} \text { if } \mu_{n}^{2}+\nu_{j}^{2}>k^{2}
\end{array}\right.
$$

The value of $B_{n}$ is set by the normalization in Eq. (9.3.30), which leads to

$$
\begin{equation*}
B_{n, j}=\sqrt{2} \text { if } j=0, \text { otherwise } B_{n, j}=2 \tag{3}
\end{equation*}
$$

Alignment of the transducer in the vertical direction means that the axial velocity everywhere on the cross section at $x=0$ is zero, except along the line $z=W / 2$. Thus, the normal velocity at $x=0$ is $V_{s}=\tilde{Q}_{0} \delta(z-W / 2)$. The corresponding velocity coefficients are indicated by Eq. (9.3.37) to be

$$
\begin{align*}
V_{n, j} & =\frac{B_{n, j} \tilde{Q}_{0}}{H} \phi_{z, j}(z=W / 2) \int_{0}^{H} V_{s} \phi_{y, n} d y \\
& =\frac{2 B_{n, j} \tilde{Q}_{0}}{(2 n-1) \pi W} \sin \left(\frac{(2 n-1) \pi}{2}\right) \cos \left(\frac{j \pi}{2}\right) \tag{4}
\end{align*}
$$

The series length is set by the largest cutoff mode numbers for both directions, which are $N_{\text {cut }}=$ floor $\left((k H / \pi+1 / 2)\right.$ and $J_{\text {cut }}=$ floor $(k W / \pi)$. For the present system parameters ( $f=20 \mathrm{kHz}, c=1480 \mathrm{~m} / \mathrm{s}, H=0.6 \mathrm{~m}, W=0.5 \mathrm{~m}$ ), these values are $N_{\mathrm{ev}}=16$, and $J_{\mathrm{ev}}=13$. We wish to construct a scan along locations that are close to the boundary, so evanescent modes must be included. Truncation of the series at $n=3 N_{\mathrm{ev}}$ and $j=3 J_{\mathrm{ev}}$ should be adequate. In principle, we could reduce these numbers as $x$ increases. However, the algorithm that follows is quite efficient, so there is no need to reduce the computational effort.

We have seen in a number of contexts that it is desirable to evaluate a modaltype series by the matrix algorithm employed in Example 9.6. Doing so enables the computation to be carried out in a vectorized manner. To use that algorithm here we must arrange the modes in a sequence described by a single subscript. Let us define an index $m=1,2, \ldots$ that is in a one-to-one correspondence to the $n, j$ indices of a transverse mode function. This can be done by incrementing $m$ as $n$ is increased with $j$ fixed, then repeating the process with $j$ increased by one, until the range of $n$ is exhausted. Because $1 \leq n \leq N$ and $0 \leq j \leq J$, the rule is

$$
\begin{equation*}
m=n+j N \tag{5}
\end{equation*}
$$

Integer arithmetic allows us to identify the $n$ and $j$ values associated with a specific $m$. If we divide the preceding by $N$ we find that $m / N \geq j$, which leads to

$$
\begin{equation*}
j=\text { floor }\left(\frac{m-\varepsilon}{N}\right), \quad n=m-j N \tag{6}
\end{equation*}
$$

where $\varepsilon$ is a very small number required to handle round-off error.
This arrangement allows us to implement the algorithm described by Eq. (4) of Example 9.6. The first step is to rewrite the modal series in a single subscript form. The number of modes is $N(J+1)$, so the series becomes

$$
P \approx \rho_{0} c \sum_{m=1}^{N(J+1)} \frac{k}{\kappa_{m}} V_{m} \Phi_{m}(y, z) e^{-i \kappa_{m} x}
$$

To convert this to matrix form let the diagonal array [ $K$ ] hold the factors in the sum that do not depend on position,

$$
\begin{equation*}
K_{m, m}=\frac{k}{\kappa_{m}} V_{m} \tag{7}
\end{equation*}
$$

The exponential terms are placed in [E], with a row holding the terms for all $m$ at fixed $x$. The number of rows is the number of $x$ locations at which the pressure will be computed. Thus,

$$
\begin{equation*}
E_{q, m}=\exp \left(-i \kappa_{m} x_{q}\right) \tag{8}
\end{equation*}
$$

The $[F]$ array holds the values of the transverse locations at selected points on a cross section. In any case a column of $[F]$ holds the $\Phi_{m}$ values for all $m$ at a designated $y$ and $z$. If we are interested in the pressure distribution over a range of $y$ locations at a fixed $z$, then $[F]$ is formed by adjoining columns for a sequence of $y$ values. Similarly, if our interest is in a range of $z$ locations at fixed $y$, then the adjoined columns correspond to the successive $z$ values. Thus,

$$
\begin{equation*}
F_{m, r}=\Phi_{m}\left(y_{r}, z\right) \text { or } F_{m, r}=\Phi_{m}\left(y, z_{r}\right) \tag{9}
\end{equation*}
$$

In either case, the pressure evaluation is

$$
\begin{equation*}
[P]=\rho_{0} c[E][K][F] \tag{10}
\end{equation*}
$$

The computational reduction achieved by this algorithm compared to one that actually accumulates sums is quite substantial. Indeed, the computations here, for which the total number of modes is 1920 and the number of points along the axial centerline was 1000 , required 0.60 s to execute on a laptop computer with an Intel I7 processor and 8 GB of RAM. Truncating the modal series at $2 N_{\text {cut }}$ and $2 J_{\text {cut }}$ reduced this time to 0.37 s .

A few details of the MATLAB program are worth reviewing. The vector $\{m\}=$ $\{12 \cdots N(J+1)\}$ consists of the single mode indices. Vectors $\{n\}$ and $\{j\}$, which hold the respective mode numbers corresponding to each element of $\{m\}$, were set according to Eq. (6). (The length of $\{n\}$ and $\{j\}$ is the same as the length of $\{m\}$.) The eigenvalues corresponding to $\{n\}$ and $\{j\}$ are placed in arrays $\{\mu\}$ and $\{\nu\}$. The $\kappa_{m}$ values also are computed in a vectorized manner. In MATLAB, this operation is kappa=conj ( $\operatorname{sqrt}\left(k^{\wedge} 2+m u .^{\wedge} 2-n u .^{\wedge} 2\right)$ ), where ". ${ }^{\wedge}$ " indicates that the square should be done element by element. (The use of the conjugate here is dictated by MATLAB's implementation of the branch cut for a square root.) The vector $\{\kappa\}$ is used to compute the successive rows of $[E]$, with $x$ held constant in each row. The evaluation of $[F]$ is done in a row-wise manner using $\{\mu\}$ and $\{\nu\}$, with the $y$ and $z$ values fixed in a column.

Figure 1 shows $|P|$ as a function of $x$ along the geometric centerline, $y=H / 2$, $z=W / 2$. The overall pressure amplitude is comparable to the result in Example 9.6. That is, the pattern of relatively broad peaks separated by narrow antinodes is like the preceding result. This is to be expected because Eq. (9.3.28) shows that the signal is a superposition of plane waves, just as it is in the two-dimensional case.


Figure 1.

Profiles of $|P|$ along a cross section appear in Fig. 2. The upper pair describe $x=0$ and the lower pair are the distributions at $x=3 \mathrm{~m}$. The vertical axis of the graphs on the right depicts the $y$ dependence because that is the way the $y$-axis is physically aligned in Example 9.6. Here too we see interference patterns


Figure 2.
resulting from superposing the contributions of numerous modes. At $x=0$ the scan along $z=W / 2$ shows the pressure on the transducer. Interference leads to fluctuations of $|P|$ on the face of the transducer, even though $V_{x}$ is constant. The scan along $y=H / 2$ shows a spike of $|P|$ in the vicinity of the transducer. This is like the behavior on the face of a translating piston, where a uniform particle velocity generates a fluctuating face pressure. Neither pattern is evident at the farther cross section because the location is many wavelengths from the transducer.

The symmetry properties of these profiles is important. The waveguide's shape is symmetric with respect to the vertical and horizontal mid-planes, that is, the planes that contain the axial centerline. The pressure does not share this symmetry, because the upper boundary is pressure-release, whereas the lower boundary is rigid. However, it is symmetric with respect to the vertical mid-plane because the walls to the left and right are rigid, and the transducer properties are the same on either side of that plane. This symmetry is manifested in Eq. (4), which indicates that the $V_{n, j}$ velocity coefficients are identically zero if $j$ is odd. As a result, all modes in the vertical direction are excited, but only the even modes for the horizontal dependence are excited. The consequence is that a scan of $|P|$ along a horizontal line is an even function with respect to the middle, whereas a scan along a vertical line shows no pattern repetition. We could have exploited this symmetry by limiting the evaluation of $|P|$ to the range $0 \leq z \leq W / 2$ and only including the even functions of $z$ in the modal series.

### 9.3.3 Circular Waveguide

Circular waveguides are commonly used in such diverse applications as HVAC and nuclear reactor cooling systems. Cylindrical coordinates, with the axial direction designated as $x$, are appropriate to this configuration. Propagation in a circular waveguide is the complement of radiation from a vibrating cylinder. Both are governed by the Helmholtz equation in cylindrical coordinates. The difference is that the waves propagate in the transverse direction in the radiation problem, whereas the propagation is axial for a waveguide. Cylindrical Bessel functions are part of the general solution of the Helmholtz equation in cylindrical coordinates. Consequently, much reference will be made to Sect.7.3.2, where the fundamental properties of these functions are described.

## Fundamental Solution for a Waveguide Mode

The pressure is taken to be a wave that propagates in the axial direction with wavenumber $\kappa$. The complex amplitude is allowed to depend on the radial distance $R$ and circumferential angle $\theta$. The $\theta$ dependence is set by the fact that $(R, \theta \pm 2 \pi, z)$ and $(R, \theta, z)$ designate the same point, so that any signal must be expressible as a Fourier series in $\theta$ with a period of $2 \pi$. We shall use a complex exponential to express the $\theta$-dependence of a term in this series. Thus, the ansatz we have identified is

$$
\begin{equation*}
p=\operatorname{Re}\left(f(R) e^{-i n \theta} e^{-i \kappa x} e^{i \omega t}\right) \tag{9.3.39}
\end{equation*}
$$

The choice of $\kappa$ as the axial wavenumber is consistent with the general analysis in Sect.9.2.1. However, the variable for the transverse dependence will be $\mu R$. This notation is opposite the usage in Sect.7.3, where the transverse wavenumber of a cylindrical wave was $\kappa$ and the axial wavenumber was $\mu$. One should bear this in mind when reference is made to prior developments. Another difference from the analysis of radiation is that the axial wavenumber was taken to be known as part of the specification of surface motion, whereas here it is a parameter to be determined.

Substitution of the assumed form of $p$ into the Helmholtz equation in cylindrical coordinates leads to

$$
\begin{equation*}
\frac{d^{2} f}{d R^{2}}+\frac{1}{R} \frac{d f}{d R}+\left(\mu^{2}-\frac{n^{2}}{R^{2}}\right) f=0 \tag{9.3.40}
\end{equation*}
$$

where

$$
\begin{equation*}
\mu^{2}=k^{2}-\kappa^{2} \tag{9.3.41}
\end{equation*}
$$

This is Bessel's equation, which we previously encountered as Eq. (7.3.8). Its fundamental solutions are the Bessel function $J_{n}(\mu R)$ and the Neumann function $N_{n}(\mu R)$, so the general solution is

$$
\begin{equation*}
f=B_{n} J_{n}(\mu R)+C_{n} N_{n}(\mu R) \tag{9.3.42}
\end{equation*}
$$

The Neumann function is singular at $R=0$, so it is discarded by setting $C_{n}=0$.
Satisfaction of the boundary condition on the cylindrical wall will be seen to require specific values of $\mu$; these are the eigenvalues. These values will depend on which azimuthal harmonic is being considered. Each $\mu$ value will lead to a different $\kappa$, so the modes are dispersive. With $n$ restricted to be positive or zero, the waveguide modes and associate transverse mode functions are given by

$$
\begin{align*}
P & =\Phi(R, \theta) e^{-i \kappa x}  \tag{9.3.43}\\
\Phi(R, \theta) & =\left\{\begin{array}{l}
B J_{n}(\mu R) e^{-i n \theta} \\
\text { or } \\
B J_{n}(\mu R) e^{i n \theta}
\end{array}\right.
\end{align*}
$$

Adding and subtracting these fundamental solutions yields sine and cosine functions for the azimuthal dependence. Thus, an alternative form of the waveguide modes is

$$
\begin{align*}
P & =\Phi(R, \theta) e^{-i \kappa x}  \tag{9.3.44}\\
\Phi(R, \theta) & =\left\{\begin{array}{l}
B J_{n}(\mu R) \cos (n \theta) \\
\text { or } \\
B J_{n}(\mu R) \sin (n \theta)
\end{array}\right.
\end{align*}
$$

The difference between Eqs. (9.3.43) and (9.3.44) is that the former describes a pair of helical waves that spiral about the $x$-axis in opposite senses, whereas the latter describes standing wave patterns around the $x$-axis. The sine and cosine functions are $90^{\circ}$ out-of-phase and therefore constitute separate solutions. In the special case where $n=0$, the two helical wave coalesce to a plane wave that propagates in the axial direction, and the sine mode vanishes.

Just as the pattern in the circumferential direction may be viewed as either oppositely traveling waves or standing waves, the same is true for the transverse direction. According to Eq. (7.3.11), the Bessel function is equivalent to the sum of both kinds of Hankel functions,

$$
\begin{equation*}
J_{n}(\mu R)=\frac{1}{2}\left[H_{n}^{(1)}(\mu R)+H_{n}^{(2)}(\mu R)\right] \tag{9.3.45}
\end{equation*}
$$

Thus, any of the descriptions of the modes may equivalently be considered to be waves that propagate inward and outward from the center. Each wave individually would be singular at $R=0$, but their singularities cancel.

In addition to finiteness, any pressure field must be continuous at $R=0$. This means that the same pressure must be obtained as a field point approaches $R=0$ along any $\theta$, which is stated as

$$
\begin{equation*}
\lim _{R \rightarrow 0} P(R, \theta, x)=g(x) \tag{9.3.46}
\end{equation*}
$$

where $g(x)$ is used to indicate that the limit may depend on $x$. Continuity of $\nabla P$ also is a fundamental requirement. Otherwise, $\nabla^{2} P$ will be infinite, which according to the Helmholtz equation corresponds to infinite $P$. As $R \rightarrow 0$, the $\bar{e}_{\theta}$ direction becomes meaningless, so we may focus on the behavior of $\partial P / \partial R$ at $R=0$. Because $\bar{e}_{R}(\theta+\pi)=-\bar{e}_{R}(\theta)$, the transverse derivative at the origin must change sign when $\theta$ is incremented by $180^{\circ}$,

$$
\begin{equation*}
\left.\frac{\partial P}{\partial R}\right|_{R=0, \theta \pm \pi}=-\left.\frac{\partial P}{\partial R}\right|_{R=0, \theta} \tag{9.3.47}
\end{equation*}
$$

Let us assess whether these conditions are satisfied by the waveguide modes. All modes have the form $\Phi_{n}(R, \theta) \exp (-i \kappa x)$ with $\Phi_{n}(R, \theta)=f_{R}(R) f_{\theta}(\theta)$. The transverse dependence is $f_{R}(R)=J_{n}(\mu R) \approx(\mu R / 2)^{n} / n!$ when $R$ is small, and $f_{\theta}(\theta)=\exp (-i n \theta)$. For $n>0, f_{R}=0$ at $R=0$, so the $g$ function is zero independent of $x$. Thus, the modes for $n \geq 1$ are continuous. For $n=0$, we have $f_{R}=1$ at $R=0$ and $f_{\theta}=1$, so the $n=0$ modes also are continuous. In regard to continuity of the gradient, for $n \geq 1$, the radial derivative at the center is $\partial \Phi_{n} / \partial R=(\mu R / 2)^{n-1} \exp (-i n \theta) /(n-1)$ ! evaluated at $R=0$. This quantity is zero for $n \geq 2$, which satisfies Eq. (9.3.47). For the $n=1$ modes, we have $\partial \Phi_{n} / \partial R=\exp (-i \theta)$ at $R=0$, which also satisfies Eq. (9.3.47). The $n=0$ modes also are such that $\partial \Phi_{0} / \partial R=0$ at $R=0$. Thus, the Bessel function modes satisfy the continuity requirements at the origin.

Figure 9.14a and b, respectively, shows the behavior of the $n=0$ and $n=1$ waveguide modes in the vicinity of $R=0$. These are contrasted by Fig. 9.14c, which shows a discontinuous pressure field described by $J_{0}(\mu R) \operatorname{Re}\left(e^{i \theta}\right)$, and by Fig. 9.14d, where $J_{1}(\mu R)$ has a discontinuous gradient at $R=0$.


Fig. 9.14 Behavior of pressure fields in the vicinity of $R=0$ : a The $n=0$ propagation mode, $P=B J_{0}(\mu R)$. b The $n=1$ propagation mode, $P=B J_{1}(\mu R) e^{-i \theta}$. c A discontinuous pressure field, $B J_{0}(\mu R) e^{-i \theta}$. d A field having a discontinuous gradient, $P=B J_{1}(\mu R)$

The radial wavenumber is set by the boundary condition at the wall, where $R=a$. A positive value of $V_{R}$ corresponds to movement into the wall, so an impedance boundary condition corresponds to

$$
\begin{equation*}
P=\rho_{0} c \zeta V_{R}=-\frac{\zeta}{i k} \frac{\partial P}{\partial R} \text { at } R=a \tag{9.3.48}
\end{equation*}
$$

When we substitute the general solution for a waveguide mode into this condition, the exponential factors cancel, as does the amplitude coefficient $B$. What remains is the characteristic equation,

$$
\begin{equation*}
J_{n}(\mu a)-\frac{i \zeta}{k a}(\mu a) J_{n}^{\prime}(\mu a)=0 \tag{9.3.49}
\end{equation*}
$$

Regardless of the nature of $\zeta$, the characteristic equation for a specified value of the circumferential number $n$ will have an infinite number of roots. These nondimensional roots of the characteristic equation are denoted as $\eta_{n, j}$. The corresponding eigenvalues are $\mu_{n, j}=\eta_{n, j} / a$. The eigenvalues might be imaginary or complex, so we shall arrange them sequentially in increasing order of their real part, $\operatorname{Re}\left(\mu_{n, 1}\right)<\operatorname{Re}\left(\mu_{n, 2}\right)<\ldots$ The manner in which the characteristic equation has been written makes it evident that the $\eta_{n, j}$ values depend only on $\zeta$ and $k a$.

Numerical methods will be required to obtain accurate roots of the characteristic equation. Few software packages offer routines that evaluate derivatives of Bessel functions. A recurrence relation in Eq. (7.3.15) gives

$$
\begin{equation*}
J_{n}^{\prime}(\mu a)=-J_{n+1}(\mu a)+\left(\frac{n}{\mu a}\right) J_{n}(\mu a) \tag{9.3.50}
\end{equation*}
$$

This expression may be incorporated as an auxiliary routine, but it is just as easy to incorporate it into the characteristic equation, which becomes

$$
\begin{equation*}
J_{n}(\mu a)-\frac{i \zeta}{k a}(\mu a)\left[-J_{n+1}(\mu a)+\left(\frac{n}{\mu a}\right) J_{n}(\mu a)\right]=0 \tag{9.3.51}
\end{equation*}
$$

The methods that are commonly used to solve transcendental equations require initial guesses of the roots $\eta_{n, j}=\mu_{n, j} a$. If $\zeta$ is infinite, zero, or purely imaginary, the equation is real. Guesses for the lower roots in that case may be obtained by graphing the characteristic equation as a function of $\mu a$. For large $\mu a$ ( $\mu a>8$ is reasonable), we may employ the asymptotic representation of $J_{n}(\mu a)$ for large arguments

$$
\begin{equation*}
J_{n}(\mu a) \rightarrow\left(\frac{2}{\pi \mu a}\right)^{1 / 2} \cos \left(\mu a-\frac{2 n+1}{4} \pi\right) \tag{9.3.52}
\end{equation*}
$$

When $\mu a$ is large, increasing $\mu a$ changes the square root factor much less than the cosine term. It follows that for large $\mu a$, the characteristic equation is approximately periodic in $\Delta(\mu a)=2 \pi$. The consequence of this property is that after we have found a large eigenvalue $\mu_{n} a \gg 1$, a good starting guess for the next value is $\mu_{n+1} a=\mu_{n} a+\pi$.

Two indices are required to uniquely specify which eigenvalue is under consideration: the circumferential harmonic number $n$ and the root number $j$, which gives the sequential placement of the root within the set of eigenvalues at that $n$. Specification of a transverse mode function requires a third index, denoted as $\alpha$. This index is used to specify the dependence on $\theta$. We will adopt the convention that if this subscript is $\alpha=1$ or $\alpha=-1$, we are using helical waves, such that

$$
\begin{equation*}
\Phi_{n, j, \alpha}=B_{n, j, \alpha} J_{n}\left(\mu_{n, j} R\right) e^{i \alpha n \theta} \tag{9.3.53}
\end{equation*}
$$

Alternatively, if we wish to use real sinusoidal terms to describe the $\theta$ dependence, we would set $\alpha=c$ or $\alpha=s$, such that

$$
\begin{align*}
& \Phi_{n, j, c}=B_{m, n, c} J_{n}\left(\mu_{n, j} R\right) \cos (n \theta)  \tag{9.3.54}\\
& \Phi_{m, n, s}=B_{n, j, s} J_{n}\left(\mu_{n, j} R\right) \sin (n \theta)
\end{align*}
$$

The discussion thus far covers the generalities. However, examination of specific impedance cases discloses issues that are not yet evident, as well as some significant behaviors.

## Rigid Walls

The model of a rigid wall serves as a reference system for discussing other configurations, as well as being useful for its own sake. Satisfaction of Eq. (9.3.49) when $\zeta$ is infinite requires that $J_{n}^{\prime}(\mu a)=0$. Thus, the eigenvalues correspond to locations at which $J_{n}(\mu a)$ has an extreme value. Figure 9.15 plots some low-order Bessel functions and shows the smallest roots $\eta_{n, j}$ for each $n$.


Fig. 9.15 Characteristic equation for a circular waveguide with rigid walls
The first few eigenvalues for a rigid-walled cylinder are listed in Table 9.1.
The axial wavenumbers are defined by Eq. (9.3.41). As was true for the previous situations, a large eigenvalue will lead to cutoff of the axial propagation. The possibilities are

Table 9.1 Roots of the characteristic equation for a rigid-walled cylindrical waveguide

| $\eta_{n, j} \equiv \mu_{n, j} a$ | $j=1$ | $j=2$ | $j=3$ | $j=4$ |
| :--- | :--- | :--- | :--- | :--- |
| $n=0$ | 0 | 3.8317 | 7.0156 | 10.1735 |
| $n=1$ | 1.8412 | 5.3314 | 8.5363 | 11.7060 |
| $n=2$ | 3.0542 | 6.7061 | 9.9695 | 13.1704 |
| $n=3$ | 4.2012 | 8.0152 | 11.3459 | 14.5858 |

$$
\kappa_{n, j}=\left\{\begin{array}{l}
{\left[k^{2}-\left(\eta_{n, j} / a\right)\right]^{1 / 2} \text { if } \eta_{n, j}<k a}  \tag{9.3.55}\\
-i\left[\left(\eta_{n, j} / a\right)^{2}-k^{2}\right]^{1 / 2} \text { if } \eta_{n, j}>k a
\end{array}\right.
$$

Because the eigenvalues are roots of a transcendental equation, it is not possible to identify a priori how many eigenvalues are below the cutoff value. Nevertheless, the tabulation indicates that the eigenvalues at a fixed root number increase with increasing $n$. Furthermore, by definition, the eigenvalues at fixed $n$ increase with increasing root number, Thus, it is sufficient to identify the root number $J$ for $n=0$ at which $\eta_{0, J}>k a$, and the circumferential harmonic $N$ at which $\eta_{1, N}>k a$. Using these as the cutoff numbers assures that all propagating modes are included. Figure 9.16 shows contours of some typical transverse mode functions. Bold contours


Fig. 9.16 Contours of the transverse mode functions for a rigid-walled circular waveguide. An $\eta_{n, j}$ value is the $j$ th root of the characteristic equation corresponding to circumferential harmonic number $n$
are zeros, which separate regions of positive and negative pressures. The patterns describe the cosine functions, $\Phi_{n, j, c}$; the sine functions would be rotated by $\pi /(2 n)$.

## Compliant Walls

Many features exhibited by the modes in the rigid-walled case are typical of those for other wall conditions. Equation (9.3.43) or the alternative form in Eq. (9.3.44) describes any case. A corollary of the occurrence of $k a$ in the characteristic equation, Eq. (9.3.49), is that it will be necessary to recompute its roots if the frequency of interest changes, for example, to perform a frequency sweep in order to determine a dispersion curve. The only exceptions are rigid or pressure-release walls.

If $\zeta$ is finite and nonzero, then neither term in the characteristic equation is zero. It follows that unless $\zeta$ is purely imaginary, the equation is complex. This means that the eigenvalues must be complex, which is a possibility we will address. The exceptional case is that where $\zeta$ is purely imaginary, corresponding to a purely reactive wall. To analyze that case, we set $\zeta=\sigma i$, with $\sigma>0$ for an inertance and $\sigma<0$ for a compliance. The corresponding form of the characteristic equation is

$$
\begin{equation*}
J_{n}(\mu a)+\sigma \frac{\mu a}{k a} J_{n}^{\prime}(\mu a)=0 \tag{9.3.56}
\end{equation*}
$$

The Bessel function is oscillatory, from which it follows that this equation has an infinite number of roots $\eta_{n, j}=\mu_{n, j} a, j=1,2, \ldots$

Let us consider the behavior of the eigenvalue for a fixed circumferential harmonic $n$ and root number $j$ as the frequency increases. The nondimensional parameter affecting the eigenvalues is $\sigma /(k a)$. The case of a pressure-release wall, which is the limit as $|\zeta| \rightarrow 0$, corresponds to $J_{n}(\mu a)=0$. In contrast, $\zeta=\infty$ gives $J_{n}^{\prime}(\mu a)=0$, which is the characteristic equation for a rigid wall. It follows that this eigenvalue will be close to that of a rigid waveguide at very low frequencies, while it will be close to the eigenvalue of a pressure-release configuration for high frequencies. A different trend is that in which $n$ and $k a$ are held fixed. By definition, increasing $j$ corresponds to increasing eigenvalue. Because $\mu a$ multiplies $\sigma$ in the characteristic equation, the higher eigenvalues will approach those for a rigid wall regardless of the size of $\sigma$. We previously encountered this trend in the analysis of a two-dimensional waveguide whose walls are reactive.

Like a two-dimensional waveguide, a circular waveguide whose wall is purely reactive might have a purely imaginary eigenvalue. To explore this possibility, we set $\mu=i \beta, \beta>0$. Rather than dealing with Bessel functions whose argument is imaginary, it is preferable to convert $J_{n}(\mu a)$ to the first kind of modified Bessel function, which is given by Eq. (7.3.33) to be

$$
\begin{equation*}
J_{n}(i \beta a)=i^{n} I_{n}(\beta a) \tag{9.3.57}
\end{equation*}
$$

Substitution of this expression into the third recurrence relation for a derivative of a Bessel function in Eq. (7.3.15) gives

$$
\begin{equation*}
J_{n}^{\prime}(i \beta a)=i^{n-1}\left[I_{n+1}(\beta a)+\frac{n}{\beta a} I_{n}(\beta a)\right] \tag{9.3.58}
\end{equation*}
$$

Substitution of these relations into the characteristic equation converts it to

$$
\begin{equation*}
I_{n}(\beta a)+\operatorname{Im}(\zeta) \frac{\beta a}{k a}\left[I_{n+1}(\beta a)+\frac{n}{\beta a} I_{n}(\beta a)\right]=0 \tag{9.3.59}
\end{equation*}
$$

One feature is immediately apparent. Each of the modified Bessel functions appearing above is positive for $\beta a>0$. It follows that a real root of this characteristic equation can exist only if $\operatorname{Im}(\zeta)<0$. In other words, a mode whose eigenvalue is purely imaginary exists only if the wall impedance is a pure compliance. This condition is consistent with the previous result for a two-dimensional waveguide. We will say that this is the ( $n, \mathrm{im}$ ) mode, so $\mu_{n, \mathrm{im}}=i \beta_{n, \mathrm{im}} a$.

To get a more detailed picture, let us consider the modal properties for $\sigma= \pm 5$ ( $\zeta= \pm 5 i$ ), $n=0$, and $k a=\pi$. The characteristic equation for real eigenvalues, Eq. (9.3.56), and an imaginary eigenvalue, Eq. (9.3.59), are depicted in Fig. 9.17. The function that is plotted for imaginary eigenvalues is the above divided by $I_{0}(\beta a)$. (This is done because $I_{n}(\beta a)$ increases exponentially with increasing $\beta a$, which would make it difficult to see the location where the plotted function is zero.) The intersection of each curve with the zero axis marks a root.


Fig. 9.17 Characteristic equation for a circular waveguide whose wall is compliant, $k a=\pi, n=0$

Each eigenvalue in the range $\mu a>3$ for $\sigma=-5$ has a close companion for positive $\sigma$. The exception is the first eigenvalue for $\sigma=5, \mu_{1} a=1.215$, whereas no real eigenvalue for $\sigma=-5$ that is that small. However, the imaginary eigenvalue for $\sigma=-5$ also has a small magnitude, which suggests that they might have some similarities. The first few eigenvalues and corresponding axial wavenumbers are listed in Table 9.2. In both reactance cases, only one mode is propagative. As $k a$ is increased, more and more modes enter the spectrum that propagates.

Table 9.2 Eigenvalues of a cylindrical waveguide with a reactive wall, $\mathrm{ka}=\mathrm{pi}$

|  |  | $j=\mathrm{im}$ | $j=1$ | $j=2$ | $j=3$ |
| :--- | :--- | :--- | :--- | :--- | :---: |
| $\zeta=-5 i$ | $\eta_{0, j} \equiv \mu_{0, j} a$ | $1.2146 i$ | 3.6657 | 6.9257 | 10.1116 |
|  | $\kappa_{0, j} a$ | 3.368 | $-1.8888 i$ | $-6.1722 i$ | $-9.6112 i$ |
| $\zeta=5 i$ | $\eta_{0, j} \equiv \mu_{0, j} a$ | - | 1.0387 | 3.9910 | 7.1044 |
|  | $\kappa_{0, j} a$ | - | 2.9649 | $-2.4614 i$ | $-6.3720 i$ |

The transverse mode functions for an inertance and a compliance are depicted in Fig. 9.18. The functions associated with real eigenvalues $\mu_{0, j}=\eta_{0, j} / a$ are like those in Fig. 9.16. This is to be expected because $\zeta= \pm 5 i$ is effectively quite close to the rigid condition. However, a finite impedance requires that $d \Phi_{n, j} / d R \neq 0$ at $R=a$, so the maxima, minima, and zeros of the $R$-dependence are shifted. The single imaginary eigenvalue, $\mu_{n, \text { im }}=i \eta_{n, \mathrm{im}} / a$, corresponds to a mode function that is reminiscent of the plane wave mode. However, the first mode for $\zeta=5 i$ also shows relatively little variation along a transverse lines, so it is too reminiscent of the plane wave mode.


Fig. 9.18 Transverse mode functions for a circular waveguide whose wall is reactive, $k a=\pi$, $n=0, \zeta=-0.2 i$

EXAMPLE 9.8 An annular waveguide consists of the region between two concentric cylinders. The outer radius is $a$, and the inner radius is $b$. Consider the case where both cylinders are rigid. Derive the characteristic equation and an expression for the transverse mode functions. Evaluate the $n=0, n=1$, and $n=2$ modes for the case where $b / a=0.4$. Also, determine the highest frequency for which all nonplanar modes evanesce.

## Significance

Coaxial pipes are used in some systems to convey fluid back and forth, especially if the surroundings are narrow. The analysis suggests how the formulation for a complete cylindrical cross section may be generalized. To some extent, it is a precursor of the analyses of a cavity, which is the topic of the next chapter.

## Solution

The transverse mode function must satisfy the rigidity condition at both walls, which are

$$
\frac{\partial \Phi}{\partial R}=0 \text { at } R=a \text { and } R=b
$$

Because $R=0$ is not a point in the fluid, both the Bessel and Neumann functions in the general solution for $R$ dependence, Eq. (9.3.42), are appropriate. A transverse mode function is the product of that function and the $n$th harmonic for the $\theta$ dependence, so the form is

$$
\begin{equation*}
\Phi=\left[B_{n} J_{n}(\mu R)+C_{n} N_{n}(\mu R)\right] e^{-i n \theta} \tag{1}
\end{equation*}
$$

The coefficients are found by making this general solution fit the boundary conditions. The matrix form of the result of these operations is

$$
\left[\begin{array}{ll}
J_{n}^{\prime}(\mu a) & N_{n}^{\prime}(\mu a)  \tag{2}\\
J_{n}^{\prime}(\mu b) & N_{n}^{\prime}(\mu b)
\end{array}\right]\left\{\begin{array}{l}
B_{n} \\
C_{n}
\end{array}\right\}=\left\{\begin{array}{l}
0 \\
0
\end{array}\right\}
$$

The characteristic equation, which results from setting the determinant of these equations to zero, is

$$
\begin{equation*}
J_{n}^{\prime}(\mu a) N_{n}^{\prime}(\mu b)-N_{n}^{\prime}(\mu a) J_{n}^{\prime}(\mu b)=0 \tag{3}
\end{equation*}
$$

The frequency does not occur in this characteristic equation, which means that the eigenvalues are independent of the frequency. Figure 1 graphs the characteristic equation as a function of $\mu a$ with $\mu b \equiv(b / a) \mu a$.


Figure 1.
The zero crossing of the curve for each $n$ is used to initiate a numerical search for the respective eigenvalues. The walls are rigid, so a plane wave mode may exist.

Thus, $n=0$ is a plane wave, for which $\mu_{0,1}=0$ and $\Phi_{0,1}=1$. The result appears in the following tabulation.

|  | $\mu_{n, j} a$ |  |  |  |
| :--- | :--- | :--- | :--- | :--- |
| $n$ | $j=1$ | $j=2$ | $j=3$ | $j=4$ |
| 0 | 0 | 5.3912 | 10.5577 | 15.7665 |
| 1 | 1.46178 | 5.6591 | 10.6833 | 15.8481 |
| 2 | 2.8424 | 6.4160 | 11.0560 | 16.0916 |

When $\mu a$ is one of the eigenvalues, only one of the scalar equations described by Eq. (2) is independent. Thus, only the ratio of the coefficients is described uniquely. When we use the second coefficient equation, which is the boundary condition at $R=a$, to evaluate $C / B$, we find that

$$
\Phi_{n, j, 1}=B_{n, j}\left[J_{n}\left(\mu_{n, j} R\right)-\frac{J_{n}^{\prime}\left(\mu_{n, j} a\right)}{N_{n}^{\prime}\left(\mu_{n, j} a\right)} N_{n}\left(\mu_{n, j} R\right)\right] e^{-i n \theta}
$$

(Recall that $e^{+i n \theta}, \cos (n \theta)$, and $\sin (n \theta)$ are equally valid descriptions of the $\theta$-dependence.) Figure 2 shows some transverse mode functions for $b / a=0.4$. The value of $B_{0, j}$ for each was selected such that the maximum of the respective function is one. If we were to let $s$ be the distance from the inner cylinder, it would not be a terrible approximation to say that these modes fit $\cos (j \pi s /(a-b))$, which are the transverse modes of a hard-walled two-dimensional waveguide.


Figure 2.

## Signal Generation

Our interest here is the signal generated by the imposition of an axial particle velocity at $x=0$. If the spatial pattern of the velocity imposed at $x=0$ does not match a specific mode, all modes are excited. Thus, we consider an arbitrary distribution, $V_{x}=\operatorname{Re}\left(V_{s}(R, \theta) e^{i \omega t}\right)$ at $x=0$. The procedure for identifying the mix of waveguide modes that is generated by this excitation follows the general outline laid out in Sect.9.3.1. The first task is to decide how to represent the $\theta$-dependence of the modes. Complex exponentials are suitable for an excitation that rotates about the $x$-axis, which is typical of noise generation mechanisms in rotating machinery. However, most excitations have a fixed distribution in the $\theta$ direction, and trigonometric functions share that property. Hence, we shall use the transverse mode functions in Eq. (9.3.44),

$$
\begin{align*}
& \Phi_{n, j, c}=B_{n, j} J_{n}\left(\mu_{n, j} R\right) \cos (n \theta)  \tag{9.3.60}\\
& \Phi_{n, j, s}=B_{n, j} J_{n}\left(\mu_{n, j} R\right) \sin (n \theta)
\end{align*}
$$

In the usual situation the range of indices is $n \geq 0, j \geq 1$. This scheme is modified if the wall is rigid by designating the first mode for $n=0$ as $j=0$ to represent the plane wave mode. In addition, if the wall is purely compliant, the imaginary eigenvalue at any $n$ is designated with the label $j=$ "im". Also, for $n=0, B_{0, j, s}$ are set to zero.

According to the general orthogonality property in Eq. (9.3.8), it must be that

$$
\begin{equation*}
\int_{0}^{a} \int_{-\pi}^{\pi} \Phi_{n, j, \alpha} \Phi_{m, \ell, \beta} R d \theta d R=\pi a^{2} \delta_{n, m} \delta_{j, \ell} \delta_{\alpha, \beta} \tag{9.3.61}
\end{equation*}
$$

The Kronecker delta term $\delta_{\alpha, \beta}$ tells us that the sine and cosine modes are orthogonal, even if their eigenvalues are the same. When the subscripts are the same, the orthogonality condition defines the normalization of the $B_{n, j, \alpha}$ coefficients. Substitution of the mode function into Eq. (9.3.60) with $\mu_{n, j}=\eta_{n, j} / a$, followed by integration over $\theta$, reduces the normalization condition to

$$
\begin{equation*}
\int_{0}^{a} \int_{-\pi}^{\pi}\left(\Phi_{n, j, \alpha}\right)^{2} R d \theta d R=\left(1+\delta_{n, 0}\right) \pi\left(B_{n, j}\right)^{2} \int_{0}^{a} J_{n}\left(\eta_{n \cdot j} \frac{R}{a}\right)^{2} R d R=\pi a^{2} \tag{9.3.62}
\end{equation*}
$$

There is no need to evaluate this integral numerically. Changing the integration variable to $\chi=R / a$ converts the equation to

$$
\begin{equation*}
\left(1+\delta_{n, 0}\right)\left(B_{n, j}\right)^{2} \mathcal{I}_{n, j}=1 \tag{9.3.63}
\end{equation*}
$$

where

$$
\begin{equation*}
\mathcal{I}_{n, j}=\int_{0}^{1} J_{n}\left(\eta_{n, j} \chi\right)^{2} \chi d \chi \tag{9.3.64}
\end{equation*}
$$

Because $\eta_{n, j}$ is a root of the characteristic equation, Eq. (9.3.56), it fits the condition for which a standard formula ${ }^{2}$ is valid. The result depends on the nature of the impedance according to

$$
\mathcal{I}_{n, j}=\left\{\begin{array}{l}
\frac{1}{2} J_{n}\left(\eta_{n, j}\right)^{2} \text { if } \zeta=0  \tag{9.3.65}\\
\frac{1}{2}\left(1-\frac{\zeta^{2} n^{2}+(k a)^{2}}{\zeta^{2} \eta_{n, j}^{2}}\right) J_{n}\left(\eta_{n, j}\right)^{2} \text { if } \zeta \neq 0
\end{array}\right.
$$

Thus, we have established that the normalizing coefficient is

$$
\begin{equation*}
B_{n, j}=\frac{1}{\left(1+\delta_{n, 0}\right)^{1 / 2}\left(\mathcal{I}_{n, j}\right)^{1 / 2}} \tag{9.3.66}
\end{equation*}
$$

The incarnation of the general modal series, Eq. (9.3.9), is

$$
\begin{equation*}
P(x, R, \theta)=\sum_{n=0}^{\infty} \sum_{j=1}^{\infty} \sum_{\alpha=c, s} P_{n, j, \alpha} \Phi_{n, j, \alpha} e^{-i \kappa_{n, j} x} \tag{9.3.67}
\end{equation*}
$$

Equation (9.3.13) describes the pressure coefficients in terms of the coefficients of a modal series for the velocity input, $V_{s}(R, \theta)$. For the case of a cylinder, Eq. (9.3.12), becomes

$$
\begin{align*}
& V_{n, j, c}=\frac{1}{\pi a^{2}} \int_{0}^{a} \int_{-\pi}^{\pi} V_{s}(R, \theta) J_{n}\left(\mu_{n, j} R\right) \cos (n \theta) R d \theta d R \\
& V_{n, j, s}=\frac{1}{\pi a^{2}} \int_{0}^{a} \int_{-\pi}^{\pi} V_{s}(R, \theta) J_{n}\left(\mu_{n, j} R\right) \sin (n \theta) R d \theta d R \tag{9.3.68}
\end{align*}
$$

The corresponding pressure coefficients are

$$
\begin{equation*}
P_{n}=\rho_{0} c \frac{k}{\kappa_{n, j}} V_{n, j, \alpha} \tag{9.3.69}
\end{equation*}
$$

Thus, the modal pressure series is

$$
\begin{align*}
P(x, R, \theta)= & \rho_{0} c \sum_{n=0}^{\infty} \sum_{j=1}^{\infty} \frac{k}{\kappa_{n, j}}\left[V_{n, j, c} \cos (n \theta)+V_{n, j, s} \sin (n \theta)\right]  \tag{9.3.70}\\
& \times B_{n, j} J_{0}\left(\mu_{n, j} R\right) e^{-i \kappa_{n, j} x}
\end{align*}
$$

[^31]The associated axial particle velocity is

$$
\begin{equation*}
V_{x}(x, R, \theta)=\sum_{n=0}^{\infty} \sum_{j=1}^{\infty}\left[V_{n, j, c} \cos (n \theta)+V_{n, j, s} \sin (n \theta)\right] B_{n, j} J_{0}\left(\mu_{n, j} R\right) e^{-i \kappa_{n, j} x} \tag{9.3.71}
\end{equation*}
$$

The overall procedure for determining the response begins with solution of the characteristic equation for the eigenvalues. The normalizing coefficients $B_{n, j}$ are evaluated with the aid of Eqs. (9.3.65) and (9.3.66). Evaluation of the integrals for the velocity coefficients in Eq. (9.3.68) might require a numerical algorithm. These items are used to form the modal series, Eq. (9.3.70). The contribution from each circumferential number $n$ may be carried out sequentially. Truncation of the modal series may be done based on the highest $n$ for which $\kappa_{n, 1}<k$ and the highest $j$ for which $\mu_{0, j}<k$, as was explained after Eq. (9.3.55).

EXAMPLE 9.9 A circular piston is mounted concentrically at one end of a very long cylindrical tube filled with air. The other end is terminated with a sand barrier that eliminates reflections. The tube is rigid, with a diameter of 100 mm , and the piston diameter is 50 mm . The piston oscillates at 40 kHz . The vibration amplitude is such that the maximum on-axis pressure would be $110 \mathrm{~dB} / / 20 \mu \mathrm{~Pa}$ if the piston were mounted in an infinite baffle in free space. Determine and graph the sound pressure on the axis of the waveguide in a range from 20 mm to 1 m .

## Significance

From an operational viewpoint, this example will illustrate some of the facets of constructing a modal series. It also will demonstrate the drastic effect of confining a sound beam.

## Solution

Let $V_{0}$ be the velocity amplitude of the piston. It is given that the piston is situated concentrically, which means that the axial velocity it imposes is the same at all $\theta$. The specific function is

$$
V_{x}(x=0)=\left\{\begin{array}{l}
V_{0}, \quad 0 \leq R<b \\
0, \quad b<R \leq a
\end{array}\right.
$$

where $b=0.025 \mathrm{~m}$ is the radius of the piston and $a=0.05 \mathrm{~m}$.
Only the $n=0$ modes are excited because $V_{x}$ is an axisymmetric distribution. The transverse mode functions of a hard-walled circular waveguide were found in Sect. 9.3.3 to be

$$
\begin{gathered}
\Phi_{0, j}=B_{0, j} J_{0}\left(\eta_{0, j} \frac{R}{a}\right) \\
J_{0}^{\prime}\left(\eta_{0, j}\right) \equiv-J_{1}\left(\eta_{0, j}\right)=0
\end{gathered}
$$

The first four roots $\eta_{0, j}$ for the $n=0$ modes are listed there in Table 9.1. For the present parameters, $c=340 \mathrm{~m} / \mathrm{s}, \omega=2.513\left(10^{5}\right) \mathrm{rad} / \mathrm{s}, a=0.05 \mathrm{~m}$, the highest mode that is not cutoff is the twelfth, for which $\eta_{0,12}=35.33$. The value of $\kappa$ for this mode and the first cutoff mode are $\kappa_{0,12}=216.95 \mathrm{~m}^{-1}$ and $\kappa_{0,13}=-213.73 i \mathrm{~m}^{-1}$. At $x=20 \mathrm{~mm}$, the exponential decay factor for the latter is approximately $1.4 \%$, which clearly demonstrates that we need not be concerned here with evanescent modes.

The scaling coefficients $B_{0, j}$ are given by Eq. (9.3.66). The specific impedance is infinite, and we are only interested in $n=0$. Therefore, we may set $J_{0}^{\prime}\left(\eta_{0, j}\right)=$ $-J_{1}\left(\eta_{0, j}\right)$, which is squared in the aforementioned equation. Hence, the coefficients are

$$
B_{0,1}=1, \quad B_{0, j}=\frac{2^{1 / 2}}{J_{1}\left(\eta_{0, j}\right)} \text { if } j>1
$$

The velocity coefficients are described in general by Eq. (9.3.68). For the above mode functions, and zero $V_{x}$ for $R>b$, that expression becomes

$$
V_{0, j}=\frac{V_{0}}{\pi a^{2}} \int_{0}^{b} \int_{-\pi}^{\pi} \Phi_{0, j} R d \theta d R=\frac{V_{0}}{\pi a^{2}} B_{0, j} \int_{0}^{b} \int_{-\pi}^{\pi} J_{0}\left(\eta_{0, j} \frac{R}{a}\right) R d \theta d R
$$

Changing the integration variable to $u=\eta_{0, j} R / a$ leads to a form that is another standard integral. ${ }^{3}$ The result is that

$$
V_{0, j}=\left\{\begin{array}{l}
V_{0} \frac{b^{2}}{a^{2}} B_{0, j}, \quad j=1 \\
2 V_{0} \frac{b}{\eta_{0, j} a} B_{0, j} J_{1}\left(\eta_{0, j} \frac{b}{a}\right), \quad j>1
\end{array}\right.
$$

The first few terms are $V_{0, j} / V_{0}=0.25,-0.3763,0.063695,0.13228, \ldots$
The value of $V_{0}$ is set by the given pressure in a free space. The maximum on-axis pressure in a sound beam is twice the plane wave value, $\max (|P|)=2 \rho_{0} c V_{0}$. The specified value of 110 dB corresponds to $\max (|P|)=8.9443 \mathrm{~Pa}$, from which we find that $V_{0}=10.961 \mathrm{~mm} / \mathrm{s}$. The Rayleigh distance is $k b^{2} / 2=0.231 \mathrm{~m}$. Hence, a scan out to $x=1 \mathrm{~m}$ extends into the farfield for the infinite baffle case.

The on-axis pressure in the waveguide is obtained by evaluating the modal series, Eq. (9.3.70) at $R=0$. Because $J_{0}(0) \equiv 1$, the modal series reduces to

$$
P(x, 0)=\sum_{j=1}^{12} P_{0, j} e^{-i \kappa_{0, j} x}, \quad P_{0, j}=\rho_{0} c \frac{k}{\kappa_{n, j}} V_{0, j}
$$

[^32]Our standard technique for computing a summation like this with $x$ fixed is to define column vector $\left\{P_{j}\right\}$ that holds the coefficients and row vector $[E(x)]$ whose elements are the complex exponentials for each $j$. To evaluate it at many $x$, we stack $[E(x)]$ for each location and compute

$$
[P]=\left[\begin{array}{c}
{\left[E\left(x_{1}\right)\right]} \\
{\left[E\left(x_{2}\right)\right]} \\
\vdots
\end{array}\right]\left\{P_{j}\right\}
$$

The result is shown in Fig. 1. The on-axis pressure for the case where the piston is mounted on an infinite baffle also is shown there. The maximum pressure in the nearfield is much less than it would be in the infinite baffle case, but the pressure in the waveguide does not show spherical decay. Both the large fluctuations over short distance and the weaker signal in the closer ranges have the same explanation. The axial signal for a baffled piston results from constructive and destructive interference between equal strength center and edge waves. In contrast, we found that waveguide modes may be represented as waves that propagate in and out in the transverse direction as they propagate downstream. There are many modes, which means that there are many waves having different amplitudes. There is no location at which all of these waves are in-phase to create a maximum.


### 9.4 Homework Exercises

Exercise 9.1 The sketch shows a conical waveguide that is driven harmonically at its left end by a piston. The right end is terminated by a damping material that eliminates reflections. It is desired to compare the pressure at point $B$ at the wall according to the Webster horn equation to the result obtained by taking the signal to be a radially symmetric wave. For a waveguide that is filled with air at frequencies ranging from 20 Hz to 5 kHz , compare the predictions of both formulations for the magnitude and phase angle of the pressure at point $B$.


Exercises 9.1 and 9.2
Exercise 9.2 The sketch shows a conical waveguide that is driven harmonically at its left end by a piston. The right end is terminated by a damping material that eliminates reflections. One model for propagation considers the signal to be a radially symmetric wave. Another approximates the radius variation as an exponential expansion, according to which the radius at point $B$, where $x=0.2 \mathrm{~m}$, is related to the end radius by $a_{B}=a_{0} \exp (0.2 b)$, where $b$ has units of $\mathrm{m}^{-1}$. Determine the magnitude and phase of the pressure of the pressure at point $B$ according to each model. The frequency varies from 10 Hz to 5 kHz and the fluid is air.
Exercise 9.3 An exponential horn is filled with air. It is 1.2 m long, with end diameters of 10 mm and 100 mm . The pressure at the large end varies harmonically at 800 Hz with an amplitude of 50 Pa . The small end is open to the atmosphere and is surrounded by a large rigid baffle that is flush with the end. (a) Determine the complex amplitudes of the pressure and particle velocity at the left end. Use the end correction on Eq. (3.5.54) for this evaluation. (b) Determine the time-averaged power radiated into the fluid domain to the left of the baffle.


Exercise 9.4 An exponential horn is driven at $x=0$ by a transducer that imposes a constant velocity amplitude $V_{0}$. The specific impedance at $x=L$ is a fixed value $\zeta$ independent of the frequency. Derive expressions for the amplitudes of the forward and backward waves as a function of $k L$ when $k>b / 2$.

Exercise 9.5 It is desired to design an exponential horn in air to connect to a piston driver. The design requirements are as follows (1): The cross-sectional area at $x=0$ should be $0.8 \mathrm{~m}^{2}$. (2) At 200 Hz , the axial particle velocity at $x=0$ should be 20 dB greater than it is at $x=L$. Equation (3) The cutoff frequency should be 100 Hz . It may be assumed that there are no reflections at the open end. Determine the growth constant $b$ and the length $L$.

Exercise 9.6 Solve Exercise 9.5 for the case where the undriven end's specific impedance is $\zeta=0.2+0.1 i$.

Exercise 9.7 The small end of an exponential horn is driven by a velocity source. The other end is joined to a very long cylindrical waveguide that is terminated by a nonreflecting material. Consequently, only waves that propagate to the right exist within that section. Derive a set of equations for the pressure amplitude in the cylindrical waveguide. Assume that the waves in the cylinder are planar.


Exercise 9.8 Use the WKB approximation, rather than the exact solution of the Webster horn equation, to carry out the analysis requested in Exercise 9.1.

Exercise 9.9 A horn's radius expands quadratically with axial distance according to $a=\alpha+\beta x^{2}$. The frequency is $\omega$, and the complex pressure amplitude at $x=0$ is $P_{0}$. (a) Derive expressions describing the WKB approximation for $P(x)$.(b) Evaluate and plot $\left|P / P_{0}\right|$ as a function of $x$ corresponding to the following parameters: $\alpha=0.25$ $\mathrm{m}, \beta=0.1 \mathrm{~m}^{-1}, f=1000 \mathrm{~Hz}, \rho_{0}=1.2 \mathrm{~kg} / \mathrm{m}^{3}, c=340 \mathrm{~m} / \mathrm{s}$.

Exercise 9.10 A seismic wave in the bottom of an ocean channel induces a wave that travels horizontally at $3200 \mathrm{~m} / \mathrm{s}$. The dominant harmonic component of this wave occurs at 12 Hz , and the amplitude of the vertical displacement in this component is 2 mm . The channel's depth is 150 m . What is the maximum pressure amplitude within the channel, and at what depth does it occur?

Exercise 9.11 A scan across the width of a two-dimensional waveguide has yielded the transverse mode function plotted below. Use the measured data to identify the specific local impedance of each wall and the eigenvalue of this mode. The fluid is air, the frequency is 400 Hz , and the width of the waveguide is 5 m .


Exercise 9.12 A porous tile is used as the bottom of a two-dimensional water channel whose height is $H$. A locally reacting model of the bottom sets the specific impedance as $\zeta=i(\beta \omega-\gamma / \omega)+\delta$. The surface of the channel may be considered to be pressure-release. (a) Derive the characteristic equation whose roots are the eigenvalues of the transverse mode functions at a specified value of $\omega$. (b) In the expression for $\zeta$, the parameter $\beta$ is proportional to a mass per unit surface area, $\gamma$ is proportional to a stiffness per unit surface area, and $\delta$ is a dissipation factor. What is the significance of the frequency being $(\gamma / \beta)^{1 / 2}$ in the situation where $\delta=0$ ? (c) For the ideal case in which $\delta=0$, derive an equation whose roots are the frequencies at which the various modes are cutoff.

Exercise 9.13 A waveguide is formed by two parallel walls having infinite extent and separated by distance $H$. One wall is well approximated as pressure-release, while the other is locally reacting with specific impedance $\zeta$. For the case where $k H=4 \pi$ and $\zeta=-2 i$, determine the eigenvalues and axial wavenumbers of the first four transverse modes. Plot each mode as a function of the transverse distance $y / H$, with each mode scaled such that $\max \left(\left|\Phi_{n}\right|\right)=1$.

Exercise 9.14 Solve Exercise 9.13 for the case where the specific impedance of the locally reacting wall is $\zeta=0.9-2 i$, and all other parameters are as specified there.

Exercise 9.15 The rigid walls of a two-dimensional waveguide are separated by distance $H$. The far end inhibits reflections. The waveguide is excited by a line source whose volume velocity per unit length is $\hat{Q}$. The source is perpendicular to the $x y$ plane at $y=H / 2$ on the cross section $x=0$. Except for the source, $v_{x}=0$ along the cross section at $x=0$. (a) Derive an expression for the pressure at an arbitrary location for any frequency $k H$. Which transverse modes contribute to the response? (b) Suppose the value of $k H$ is very close to $2 \pi$. What is the nature of the pressure distribution along any cross section at large $k x$ ?

Exercise 9.16 A two-dimensional horizontal waveguide has a rigid upper and lower walls. The velocity normal to the vertical wall at $x=0$ is a translational oscillation over the lower half, while the upper half of the wall is stationary. Thus, the boundary condition at $x=0$ is $v_{x}=v_{0} \cos (\omega t)[h(y)-h(y-H / 2)]$. The fluid is oil, $\rho_{0}=930 \mathrm{~kg} / \mathrm{m}^{3}, c=1320 \mathrm{~m} / \mathrm{s}$, the height $H=5 \mathrm{~m}$, and $\omega=25000 \mathrm{rad} / \mathrm{s}$. (a) Determine the value of $\left|v_{0}\right|$ if it is observed that the sound pressure level is $195 \mathrm{~dB} / / 1$ $\mu \mathrm{Pa}$ at a point at $y=H / 2$ on the cross section at $x=8 \mathrm{~m}$. (b) Repeat the analysis in Part (a) for the case where the sound pressure level is 195 dB at $y=H / 4, x=8 \mathrm{~m}$.

Exercise 9.17 The walls at $y=0$ and $y=H$ of a rectangular waveguide are rigid, while the walls at $z=0$ and $z=W$ are pressure-release. The end $x=0$ is rigid, except that a point source is situated at the corner where $x=y=z=0$. The termination at the far end of the waveguide prevents reflection of waves. Derive an expression for the pressure at an arbitrary location.

Exercise 9.18 A signal propagates in the horizontal $x$ direction within a long waveguide whose walls at $y=0, y=H, z=0$, and $z=W$ are rigid. The
termination at the far end inhibits reflections. The signal is generated by a harmonic vibration at the end $x=0$ whose spatial pattern is a mix of the plane wave mode and the $(2,2)$ mode, that is, $v_{x}(x=0)=\left[v_{0}+v_{2} \cos (2 \pi y / H) \cos (2 \pi z / W)\right] \cos (\omega t)$. The fluid is air, with $\omega=500 \pi \mathrm{rad} / \mathrm{s}, H=3 \mathrm{~m}, W=4 \mathrm{~m}$, and $v_{0}=v_{2}=400$ $\mathrm{mm} / \mathrm{s}$. (a) Evaluate $|P|$ as a function of $x$ along the center line, $y=H / 2, z=W / 2$ for $0 \leq x \leq 10 \mathrm{~m}$. (b) Evaluate $|P|$ as a function of $y$ along the transverse line $z=W / 2$ at $x=1 \mathrm{~m}$ and $x=10 \mathrm{~m}$.

Exercise 9.19 The walls at $y=0, y=H$, and $z=W$ of a rectangular waveguide are rigid, and the wall at $z=0$ is pressure-release. The excitation is a uniform velocity distribution over one-quarter of the end $x=0$, specifically $v_{x}=v_{0} \cos (\omega t)$ for $0<y<H / 2$ and $0<z<W / 2$. The aspect ratio is $W / H=2.5$, and $k H=10$. Derive an expression for the pressure at an arbitrary location knowing that the termination at the far end of the waveguide prevents reflection of waves. Then, evaluate this result for the pressure along the lines $y=H / 2$ and $z=W / 2$ at the cross section $x=20 \pi / k$.

Exercise 9.20 The sketch shows a two-dimensional model of a rectangular waveguide in which a layer of liquid having density $\rho_{2}$ and sound speed $c_{2}$ overlays a denser liquid whose properties are $\rho_{1}$ and $c_{1}$. The bottom is rigid, whereas air is above the upper liquid. Consequently, that interface may be considered to be pressure-release from the viewpoint of the liquid. Vibration of the vertical wall at $x=0$ generates waves that propagate to the right. The far end attenuates these waves, so that there are no return reflections. (a) Derive the characteristic equation governing horizontal propagation modes. (b) Derive an expression for the transverse mode function corresponding to an eigenvalue that would be obtained by solving the characteristic equation. (c) Prove that if $\rho_{2}=\rho_{1}$ and $c_{2}=c_{1}$, then the modal properties reduce to those of two-dimensional waveguide of height $H$ whose walls are rigid and pressure-release. Hint: Let $\mu$ and $\eta$ denote the transverse wavenumber of $\Phi$ for $0<y<H / 2$ and $H / 2<y<H$, respectively. The pressure and vertical particle velocity derived from a mode function must be continuous. In addition to leading to the conditions that $\Phi$ and $d \Phi / d y$ be continuous at $y=H / 2$, this requirement tells us that the axial wavenumber $\kappa$ must be the same for both layers. Thus, it must be that $\kappa=\left(\omega^{2} / c_{1}^{2}-\mu^{2}\right)^{1 / 2}=\left(\omega^{2} / c_{2}^{2}-\eta^{2}\right)^{1 / 2}$.


Exercise 9.21 The wall of a cylindrical waveguide is locally reacting, with specific impedance $\zeta=4 i$. Determine the eigenvalues of the first four axisymmetric transverse modes. Plot dependence of these mode functions on the transverse distance $R$. Carry out the analysis for $k a=1$ and $k a=10$.

Exercise 9.22 The wall of a cylindrical waveguide is locally reacting, with specific impedance $\zeta=4 i$. Determine and graph the phase and group veocities of the first four modes from $k a=15$ down to the cutoff frequency of each mode.

Exercise 9.23 A circular waveguide whose wall is rigid is drive by a vibrating plate at $x=0$. The spatial pattern of the vibration matches the fundamental nonplanar mode of the waveguide, that is, $v_{x}=v_{0} \Phi_{0,1} \cos (\omega t)$. Determine and graph the nondimensional mean-squared pressure $\left(p^{2}\right)_{\mathrm{av}} /\left(\rho_{0} c v_{0}\right)^{2}$ and nondimensional meansquared particle velocity $\left(v_{x}^{2}\right)_{\mathrm{av}} / v_{0}^{2}$, averaged over any cross section. Perform the analysis for frequencies ranging from cutoff to $k a=10$. Compare these dependencies to that of the power transported in that mode.

Exercise 9.24 A cylindrical waveguide whose wall is rigid is excited by a point source situated on the centerline at $x=0$. Determine the pressure amplitude on-axis at $k x=40$ as a function of frequency in the range $0.1 \leq k a \leq 16$.

Exercise 9.25 A cylindrical waveguide whose wall is rigid is excited by a point source situated at $x=0$. The location of the point source is $R=a / 2$ from the centerline. Derive an expression for the pressure field generated by the source.

Exercise 9.26 The region between the inner and outer cylinders is an annular waveguide. A disturbance interior to the inner cylinder induces a harmonic displacement wave that propagates along the inner wall. This wave propagates downstream at sound speed $c_{w}>c$. In other words, the transverse velocity at $R=b$ is $v_{R}=V \sin \left(\omega t-\omega x / c_{\mathrm{w}}\right)$. The outer cylinder is rigid. (a) Derive an expression for the pressure wave that is generated in the annular region between the cylindrical walls. Does the analysis indicate that there are values of $\omega$ that cause a resonance? (b) For the case where $c_{w}=3 c$, evaluate the pressure at the outer wall as a function of $k a<20$.


Exercise 9.26

Exercise 9.27 A circular disk is the closure of a very long cylindrical waveguide. The disk is mounted on a shaft that rotates at angular speed $\Omega$. The shaft's centerline is collinear with the centerline of the waveguide. Because of a manufacturing error, the disk is not correctly aligned. Rather, the normal to its surface forms a small constant angle $\Delta$ relative to the shaft. The sketch is a side view of the configuration when the plane containing both centerlines is vertical. The effect of the misalignment is to cause a fluctuation in the $x$ coordinate of the boundary. Consider two points at fixed
$R$ and $\theta$. One is on the nominal end, $x=0$, and the other is on the disk. This distance oscillates as a consequence of the rotation, and there is an axial velocity at $(R, \theta)$ at the end. This velocity is $v_{x}=\Delta \Omega R \sin (\Omega t-\theta)$. The cylindrical wall is rigid, and the far end is terminated in a manner that inhibits reflections. Derive an expression for the pressure within the waveguide. Which transverse modes are excited?


Exercise 9.28 A cylindrical waveguide, whose wall is rigid, is excited at the end $x=$ 0 . Waves that reach the far end are fully absorbed. The excitation at $x=0$ consists of a harmonic axial velocity that is uniformly distributed over two nonadjacent quadrants. The velocity in one quadrant is $180^{\circ}$ out-of-phase from the other, and the velocity in the intervening quadrants is zero. What transverse modes are excited? Derive an expression for the pressure at an arbitrary field point.

Exercise 9.29 A different type of waveguide that features cylindrical waves occurs in an ocean channel. The cylinder in the sketch extends from the ocean floor up to the free surface at the atmosphere. The acoustic domain here is $R>a, 0<z<H$, $-\pi<\theta<\pi$. The surface of the cylinder undergoes a vibration in the transverse direction, $\bar{e}_{R}$. (One context in which this configuration might arise is a pipe that brings oil to the surface from a drill hole.) Waves propagate away from the cylinder, as they do for an infinitely long cylinder whose surface vibrates. The fundamental difference is the finite extent of the axial direction. To avoid unnecessary complications, consider an axisymmetric pressure field $p(R, z, t)$. (a) Determine the transverse mode functions for this system. (Note that propagation is in the sense of increasing $R$, so the transverse modes are functions of z.) Do these modes have a cutoff frequency? (b) Describe the orthogonality and normalization conditions for the mode functions.
(c) Describe the modal series for the pressure corresponding to the general situation where the vibration at $R=a$ is a harmonic, with an arbitrary dependence on $z$, $v_{R}=\operatorname{Re}(f(z) \exp (i \omega t))$. Derive an expression for the pressure field generated by this vibration.


Exercise 9.29

## Chapter 10 <br> Modal Analysis of Enclosures

To say that a region containing an ideal fluid is a "cavity" is to imply that it is a void in some solid region. The synonymous term "enclosure" is more descriptive, in the sense that it conveys the notion that the domain is surrounded by a solid boundary. A tube terminated at both ends is the simplest configuration to analyze. Sections 2.5.3 and 3.2.2 addressed planar waves in such a system, but our concern here will extend to higher order waveguide modes. As is true for planar waves, closure results in reflection of the incident waves, thereby setting up a field that features standing waves in all directions. The developments that follow will primarily deal with regular geometries. The chapter will close by developing an approximate method for irregular cavities. The closure will be an approximate method for cavities and elastic structures.

### 10.1 Fundamental Issues

### 10.1.1 Wall-Induced Signals

In our study of waveguides in the previous chapter, the fundamental premise was that the far end is terminated in a manner that prevents reflections. We now remove that restriction. Figure 10.1 depicts a rectangular cavity. The length in the $x$ direction is $L$, and the height in the $y$ direction is $H$. To avoid unnecessary complications, we begin with a situation in which there is no spatial variation in the $z$ direction. Thus, the frequency-domain response will have the form $p=\operatorname{Re}\left(P(x, y) e^{i \omega t}\right)$.

Let us consider the situation where a known normal velocity distribution $V_{s}(y)$ exists on the face at $x=0$, and the walls at $y=0, y=H$, and $x=L$ are rigid. We know that if $L$ were infinite, the imposed velocity at $x=0$ would generate a set of waveguide modes that propagate in the direction of increasing $x$. The presence

Fig. 10.1 Coordinate
systems for a
two-dimensional cavity. An excitation acts at $x=0$ and the other walls are rigid

of a wall at finite $L$, does not alter this fact, but it does result in reflections that are manifested as modes that propagate in the direction of decreasing $x$. The $y$ dependence of both sets of modes must be the same set of transverse mode functions $\Phi_{n}(y)$. The transverse mode functions for a hard-walled two-dimensional waveguide are

$$
\begin{equation*}
\Phi_{n}(y)=B_{n} \cos \left(\frac{n \pi y}{H}\right) \tag{10.1.1}
\end{equation*}
$$

where normalization gives $B_{n}=1$ if $n=0$ and $\sqrt{2}$ otherwise. The axial wavenumber for both the forward and backward propagating modes must be $\kappa_{n}$, but we do not know a priori the modal amplitudes for either set of modes. Hence, the appropriate modal series is

$$
\begin{align*}
P & =\sum_{n=0}^{\infty}\left[D_{n, 1} e^{-i \kappa_{n} x}+D_{n, 2} e^{+i \kappa_{n} x}\right] B_{n} \cos \left(\frac{n \pi y}{H}\right) \\
& =\sum_{n=0}^{\infty}\left[D_{n, c} \cos \left(\kappa_{n} x\right)+D_{n, s} \sin \left(\kappa_{n} x\right)\right] B_{n} \cos \left(\frac{n \pi y}{H}\right) \tag{10.1.2}
\end{align*}
$$

where the representation of the $x$ dependence as sinusoidal functions makes it easier to satisfy boundary conditions.

As was noted, this representation satisfies the rigidity conditions at $y=0$ and $y=H$, and each term in the series individually satisfies the Helmholtz equation because the $\kappa_{n}$ wavenumbers are those in Eq. (9.2.10). It remains to satisfy the conditions at $x=0$ and $x=L$, which are

$$
\begin{gather*}
\left.\frac{\partial P}{\partial x}\right|_{x=0}=-i \omega \rho_{0} V_{s}  \tag{10.1.3}\\
\left.\frac{\partial P}{\partial x}\right|_{x=L}=0
\end{gather*}
$$

The second condition is satisfied by each term in the modal series if

$$
\begin{equation*}
D_{n, s} \cos \left(\kappa_{n} L\right)=D_{n, c} \sin \left(\kappa_{n} L\right) \tag{10.1.4}
\end{equation*}
$$

We use this relation to eliminate $D_{n, s}$ in Eq. (10.1.2) and then define a new coefficient $D_{n} \equiv D_{n, c} / \cos \left(\kappa_{n} L\right)$. This converts the series to

$$
\begin{equation*}
P=\sum_{n=0}^{\infty} D_{n} B_{n} \cos \left(\kappa_{n}(L-x)\right) \cos \left(\frac{n \pi y}{H}\right) \tag{10.1.5}
\end{equation*}
$$

The $D_{n}$ coefficients are found by satisfying the boundary condition at $x=0$. To that end, $V_{s}(y)$ is expanded in a series of the transverse mode functions. Orthogonality of those functions yields

$$
\begin{equation*}
V_{s}(y)=\sum_{n=0}^{\infty} V_{n} \cos \left(\frac{n \pi y}{H}\right), \quad V_{n}=\frac{1}{H} \int_{0}^{H} V_{s} B_{n} \cos \left(\frac{n \pi y}{H}\right) d y \tag{10.1.6}
\end{equation*}
$$

This series and Eq. (10.1.5) are substituted into the first of Eq.(10.1.3). Matching the coefficients of each mode leads to

$$
\begin{equation*}
P=-i \omega \rho_{0} \sum_{n=0}^{\infty} \frac{B_{n}}{\kappa_{j} \sin \left(\kappa_{n} L\right)} V_{n} \cos \left(\kappa_{n}(L-x)\right) \cos \left(\frac{n \pi y}{H}\right) \tag{10.1.7}
\end{equation*}
$$

This representation of the response to excitation on a boundary is quite similar to the result of the analysis in Sect. 9.2.5. One difference is that the dependence on $x$ here is a superposition of standing waves, rather than propagating waves. Each series term describing the spatial dependence is a forced cavity mode. A new feature contained in Eq. (10.1.7) is the possibility of a singularity when any of the axial wavenumbers are such that $\sin \left(\kappa_{n} L\right)=0$. Because $\kappa_{n}$ depends on $k$, these singularities correspond to resonant frequencies. We shall return to this feature after we consider a different type of excitation.

### 10.1.2 Source Excitation

Problems pertaining to the response within an enclosure that is induced by motion of a wall are particularly relevant to moving vehicles. However, if the enclosure is a room, the excitation of interest is likely to be a source, such as a person or a musical instrument. Suppose the cavity in question is that in Fig. 10.1, with all walls now taken to be rigid. Because there is no wall that generates propagating modes, there is no conceptual difference between the $x$ and $y$ directions. If we consider the former to be the axial direction, the transverse mode functions are $B_{n} \cos (n \pi y / H)$, whereas taking the $y$ direction to be axial leads to transverse functions that are $B_{j} \cos (j \pi x / L)$. It is not reasonable to favor either direction, so we use one of each function to represent $P$. There is no limit to the values of $n$ and $j$, so it seems reasonable to try an ansatz for the pressure that is the sum of all possible combinations of the transverse mode functions for each direction. That is, let us try

$$
\begin{equation*}
P=\sum_{j=0}^{\infty} \sum_{n=0}^{\infty} B_{j} B_{n} D_{j, n} \cos \left(\frac{j \pi x}{L}\right) \cos \left(\frac{n \pi y}{H}\right) \tag{10.1.8}
\end{equation*}
$$

This expression has the appearance of a Fourier cosine series in each direction, but we soon will find a more profound interpretation.

The above representation clearly satisfies all rigid wall conditions, but there is a fundamental difficulty-none of the terms in the double summation individually satisfy the Helmholtz equation at an arbitrary frequency. Furthermore, there is no indication of the role of a source. Suppose a line source is situated at an arbitrary location $\left(x_{0}, y_{0}\right)$. It was shown as part of the definition of Green's functions, Sect. 6.4.2, that the field of a harmonic point source is governed by an inhomogeneous Helmholtz equation, specifically Eq. (6.4.22). Thus, the pressure must satisfy

$$
\begin{equation*}
\frac{\partial^{2} P}{\partial x^{2}}+\frac{\partial^{2} P}{\partial y^{2}}+k^{2} P=-i \omega \rho_{0} \hat{Q} \delta\left(x-x_{0}\right) \delta\left(y-y_{0}\right) \tag{10.1.9}
\end{equation*}
$$

The terms in Eq. (10.1.8) will not satisfy this equation individually, but they might do so collectively. Conceptually, we want to find the set of coefficients $D_{j, n}$ for which the discrete, doubly infinite summation satisfies the field equation at a continuum of points within the cavity. Contemplating matching these different types of infinities is bewildering, so let us look at the matching process from a linear algebra perspective. Suppose we were to substitute Eq. (10.1.8) into the inhomogeneous wave equation, with arbitrary values assigned to the $D_{j, n}$ coefficients. The result would be a difference $\Delta(x, y)$ between the left and right side of the equations. Rather than setting $\Delta$ to zero at each location, we invoke the interpretation of the series as a sum of unit vectors in a linear functional space (the product of cosines) multiplied by component lengths (the $D_{j, n}$ values). Because all possible solutions of the cavity with rigid walls lie in this linear space, we can assert that the correct set of coefficients is that which causes $\Delta$ to be orthogonal to all directions in the linear space. The pressure functions are defined in the domain $0<x<L, 0<y<H$, so orthogonality constitutes an inner product over this space. The functions in the pressure series terms are sinusoidal, which means that the orthogonality property is that of a Fourier series. Hence, a function $F(x, y)$ is orthogonal to the $(j, n)$ term if

$$
\begin{equation*}
\int_{0}^{H} \int_{0}^{L} F(x, y)\left[B_{j} \cos \left(\frac{j \pi x}{L}\right)\right]\left[B_{n} \cos \left(\frac{n \pi y}{H}\right)\right] d x d y=0 \tag{10.1.10}
\end{equation*}
$$

To implement this concept, we substitute the pressure series into the inhomogeneous wave equation, multiply both sides by a selected function, and then integrate over the domain of the cavity. It is convenient to change the summation indices to $r$ and $s$, so that we may multiply the series by the $(j, n)$ function. These operations lead to

$$
\begin{align*}
\sum_{r=0}^{\infty} \sum_{s=0}^{\infty} \int_{0}^{H} \int_{0}^{L}\left[B_{r} B_{s} D_{r, s} \cos \right. & \left.\left(\frac{r \pi x}{L}\right) \cos \left(\frac{s \pi y}{H}\right)\right]\left[k^{2}-\left(\frac{r \pi}{L}\right)^{2}-\left(\frac{s \pi}{H}\right)^{2}\right] \\
& \times B_{j} B_{n} \cos \left(\frac{j \pi x}{L}\right) \cos \left(\frac{n \pi y}{H}\right) d x d y \\
& =-i \omega \rho_{0} \hat{Q} \int_{0}^{H} \int_{0}^{L} \delta\left(x-x_{0}\right) \delta\left(y-y_{0}\right) \\
& B_{j} b_{n} \cos \left(\frac{j \pi x}{L}\right) \cos \left(\frac{n \pi y}{H}\right) d x d y \tag{10.1.11}
\end{align*}
$$

On the left side, we invoke the orthogonality of the functions in each direction. Thus, only the term for which $r=j$ and $s=n$ is nonzero. In that case, inclusion of the $B_{m}$ in the integrals over $x$ and $y$, respectively, gives $L$ and $H$ factors. On the right side, we have the integral property of a Dirac delta function, which evaluates the analytical part of the integrand at the location where the delta occurs. The result is

$$
\begin{equation*}
L H\left[k^{2}-\left(\chi_{j, n}\right)^{2}\right] D_{j, n}=-i \omega \rho_{0} \hat{Q} B_{j} B_{n} \cos \left(\frac{j \pi x_{0}}{L}\right) \cos \left(\frac{n \pi y_{0}}{H}\right) \tag{10.1.12}
\end{equation*}
$$

where

$$
\begin{equation*}
\chi_{j, n}=\left[\left(\frac{j \pi}{L}\right)^{2}+\left(\frac{n \pi}{H}\right)^{2}\right]^{1 / 2} \tag{10.1.13}
\end{equation*}
$$

Substitution of the resulting expression for $D_{j, n}$ into Eq. (10.1.8) leads to

$$
\begin{align*}
P= & -i \omega \rho_{0} \frac{\hat{Q}}{L H} \sum_{j=0}^{\infty} \sum_{n=0}^{\infty} \frac{\left(B_{j} B_{n}\right)^{2}}{k^{2}-\left(\chi_{j, n}\right)^{2}} \cos \left(\frac{j \pi x_{0}}{L}\right) \\
& \cos \left(\frac{n \pi y_{0}}{H}\right) \cos \left(\frac{j \pi x}{L}\right) \cos \left(\frac{n \pi y}{H}\right) \tag{10.1.14}
\end{align*}
$$

This is a remarkable result because it is the response of the cavity to a point source at an arbitrary location. In other words, it is Green's function for that domain. Observe in this regard that the expression is reciprocal, as is evident by the fact that swapping $x$ and $x_{0}$, or $y$ and $y_{0}$ does not alter the result.

Thus, we have identified alternative sets of functions that describe the field within the cavity. The $\Phi_{n}$ functions in the previous section were derived from the propagation modes of a waveguide. Changing the excitation changes these functions. We shall refer to them as forced cavity modes. These attributes are contrasted by the functions from which Eq. (10.1.14) is constructed. We shall denote them as $\Psi_{j, n}$ and refer to them as natural cavity modes, although we often will omit "natural." These functions are independent on the nature of the excitation and therefore represent intrinsic properties of the system. A specific set of wall conditions leads to a specific set of $\Psi_{j, h}$. The functions appearing above are the two-dimensional functions for a hardwalled cavity,

$$
\begin{equation*}
\Psi_{j, n}=\cos \left(\frac{j \pi x}{L}\right) \cos \left(\frac{n \pi y}{H}\right) \tag{10.1.15}
\end{equation*}
$$

Let us substitute this function into the (homogeneous) Helmholtz equation. Doing so leads to recognition that any $\Psi_{j, n}$ is a solution when the acoustic wavenumber is $k=\chi_{j, n}$. These observations are equivalent to saying that $\chi_{j, n}$ are the eigenvalues and $\Psi_{j, n}$ are the eigenfunctions of the cavity. The natural frequencies are $\omega_{j, n}=c \chi_{j, n}$. It now is clear why Eq. (10.1.14) is singular if $k$ equals any $\chi_{j, n}$, or equivalently, when $\omega=\left(\omega_{\text {nat }}\right)_{j, n}$. Excitation of any linear system at its natural frequency leads to resonance. This phenomenon occurs because all internal actions balance to sustain a free response, so there is no remaining ability to resist an excitation. The only exception to this statement occurs if the excitation is such that the mode is not excited. For example, suppose that $k=\chi_{2,3}(j=2, n=3)$ in Eq. (10.1.14), but $x_{0}=L / 4$. Then, $\cos \left(j \pi x_{0} / L\right)=0$. Mathematically, this gives a term in the pressure series in which zero is divided by zero. The reality is that a true singularity is unlikely because it requires that $k$ exactly equals $\chi_{2,3}$. Furthermore, the analysis excluded dissipation effects that ameliorate resonances. Another aspect to consider is a property identified in Sect. 2.5.3, specifically, that resonance in a linearized analysis results in progressive growth of the transient response. As the response grows, nonlinear effects become increasingly important. Thus, the singularities at resonance merely flag exceptional situations.

An interesting aspect results from rewriting the pressure series, Eq.(10.1.14), in terms of the mode function notation, rather than explicit harmonic functions. Doing so gives

$$
\begin{equation*}
P=-i \omega \rho_{0} \frac{\hat{Q}}{L H} \sum_{j=0}^{\infty} \sum_{n=0}^{\infty} \frac{\Psi_{j, n}\left(x_{0}, y_{0}\right) \Psi_{j, n}(x . y)}{k^{2}-\chi_{j, n}^{2}} \tag{10.1.16}
\end{equation*}
$$

This form is descriptive of Green's function for any two-dimensional cavity, with the mode functions appropriate to that domain. Furthermore, three-dimensional situations are described by altering the summation to extend over three indices for the three directions. We will see how this result is obtained in Sect. 10.2.1.

The coexistence of forced and natural cavity modes, which result in different modal series, might cause some degree of confusion, so let us contrast their essential features. The $\Phi_{n}$ in Eq. (10.1.7) represent the transverse dependence of waveguide modes that propagate in opposite senses at any frequency. The $\Psi_{j, n}$ represent the pressure fields that can exist within a closed region without excitation only at the natural frequencies. Some individuals refer to natural cavity modes as standing wave modes. Perhaps the most important distinction is that natural modes exist for any enclosed domain, whereas a series of propagation modes obviously assumes that the picture of a waveguide closed at one end is appropriate. For example, although the analysis would not be simple, it is possible to find the cavity modes for an Lshaped room. However, the L-shape does not suggest a conceptual view in which waves propagate back and forth in one direction. Indeed, the direction that might be considered to be axial in one leg of the L-shape would correspondingly be viewed as the transverse direction for the other leg.

Resonance of any dissipationless system is manifested as a singularity at a natural frequency. Regardless of whether the field within a cavity is described as a series of forced or natural modes, resonances will be manifested as an exceptionally large value of a modal coefficient. Resonance occurs in Eq. (10.1.7), which is the forced mode series, when $\sin \left(\kappa_{n} L\right)$ in the denominator is zero for a specific $n$. This occurs when $\kappa_{n} L=j \pi, j=0,1, \ldots$ The axial wavenumber $\kappa_{n}$ for a waveguide mode is related to the eigenvalue $\mu_{n}$ of the transverse mode function by $\kappa_{n}^{2}=k^{2}-\mu_{n}^{2}$, and $\mu_{n}=$ $n \pi / H$ for a hard-walled waveguide. Hence, the resonances occur when $n^{2} \pi^{2} / H^{2}+$ $j^{2} \pi^{2} / L^{2}=k^{2}$. This condition is identical to the resonance condition $k=\chi_{j, n}$ for the natural cavity mode series in Eq. (10.1.14).

Each type of cavity mode has situations for which it is most useful. A series of forced modes is well suited to situations where the excitation originates on the boundary. However, a forced mode series cannot be implemented directly if the cavity is excited by one or more sources within the fluid. This is so because these modes satisfy the homogeneous Helmholtz equation. We will see in Example 10.4 that in some cases it is possible to use a forced mode analysis to determine the effect of a source. However, the situation addressed there is rather special, and its extension to more general systems would be quite difficult.

The utility of a natural cavity mode formulation is quite opposite. These functions are solutions of the Helmholtz equation that satisfy passive (that is, homogeneous) boundary conditions, so they cannot be used directly to satisfy an inhomogeneous boundary condition associated with an excitation. On the other hand, such a formulation readily handles cases of source excitation. ${ }^{1}$

### 10.2 Frequency-Domain Analysis Using Forced Cavity Modes

The developments in this section assume familiarity with the properties of propagation modes for waveguides. The various derivations and results in Chap. 9 are not repeated here.

### 10.2.1 Rectangular Enclosures

A cavity in the shape of a rectangular box, with one wall that undergoes a specified motion, is relatively uncomplicated to analyze, yet it exhibits the same phenomena as

[^33]any other cavity. The direction normal to the vibrating wall is designated as the axial direction $x$, with the axial length designated as $L$. The vibrating wall is defined to be situated at $x=0$. The velocity distribution at this boundary may be arbitrary, so it is required that $V_{x}=V_{s}(y, z)$ at $x=0$. The sidewalls are the planes $y=0, y=H$, $z=0$, and $z=W$. Conceptually, the signal within this cavity consists of modes that are radiated from the vibrating wall and propagate in the direction of increasing $x$. When they arrive at $x=L$, they are reflected, which results in waves that travel back to $x=0$, where they are reflected, ad infinitum. The waves that propagate forward and back in the axial direction are the propagation modes evaluated in Sect.9.3.2. The amplitudes of the forward and backward modes are quantities to be determined.

All walls other than the vibrating one are taken to be locally reacting, which includes the rigid and pressure-release conditions. The boundary conditions on the sidewalls are the same as those posed in Eq. (9.3.19) for a rectangular waveguide. Thus, the transverse mode functions are as described in Eqs. (9.3.25) and (9.3.26). For a rectangular waveguide, these functions are the product of the transverse modes for two-dimensional waveguides in each direction, specifically

$$
\begin{equation*}
\Phi_{j, n}=B_{j, n} \phi_{y, j}(y) \phi_{z, n}(z) \tag{10.2.1}
\end{equation*}
$$

Eigenvalues associated with $\Phi_{j, n}$ are $\mu_{j}$ for the $y$ functions, and $\nu_{n}$ for the $z$ functions. It is essential that the modal series account for the effect of all modes. A plane wave mode can exist in the $y$ direction if the walls at $y=0$ and $y=H$ are rigid. In that case, the sum over the modes for the $y$ direction would begin with $j=0$, with the first eigenvalue $\mu_{0}=0$. Another possibility is that $y=0$ and $y=H$ are purely compliant. In that case, there would be a purely imaginary eigenvalue. This mode can be assigned to $j=0$, so that $\mu_{0}=\operatorname{Im}\left(\mu_{\mathrm{im}}\right) i$. In any other case, the modal sum would begin with $j=1$. The same considerations apply to the $\phi_{z, n}(z)$ functions, whose eigenvalues are $\nu_{n}$. A further consideration is that the scaling factors $B_{j, n}$ are required to have been set according to the normalization rule stated in Eq. (9.3.29).

A modal series consists of sums in which the transverse functions multiply complex exponentials representing forward and backward propagation. The amplitudes of these exponentials are denoted as $\alpha_{j, n}$ and $\beta_{j, n}$, respectively, so the ansatz is

$$
\begin{equation*}
P=\sum_{n=0}^{\infty} \sum_{j=0}^{\infty}\left[\alpha_{j, n} e^{-i \kappa_{j, n} x}+\beta_{j, n} e^{+i \kappa_{j, n} x}\right] \Phi_{j, n}(y, z) \tag{10.2.2}
\end{equation*}
$$

The axial wavenumbers $\kappa_{j, n}$ must be such that each term in the series individually satisfies the Helmholtz equation, which means that

$$
\begin{equation*}
\kappa_{j, n}=\left(k^{2}-\mu_{j}^{2}-\nu_{n}^{2}\right)^{1 / 2} \tag{10.2.3}
\end{equation*}
$$

The modal series accounts for forward and backward waves in the $x$ directions, so we may continue the previous practice of placing $\kappa_{j, n}$ in the fourth quadrant of the complex plane.

Because the transverse mode functions satisfy the boundary conditions at the sidewalls, it only remains to make the particle velocity $V_{x}$ at $x=0$ match the imposed distribution $V_{s}(y, z)$ and to make $P / V_{x}$ be consistent with the locally reacting model at $x=L$. These conditions are

$$
\begin{align*}
\left.\frac{\partial P}{\partial x}\right|_{x=0} & =-i \omega \rho_{0} V_{s}(y, z)  \tag{10.2.4}\\
\left.P\right|_{x=L} & =-\left.\frac{\zeta_{L}}{i k} \frac{\partial P}{\partial x}\right|_{x=L}
\end{align*}
$$

Substitution of the modal series into these boundary conditions gives

$$
\begin{gather*}
\sum_{n=0}^{\infty} \sum_{j=0}^{\infty} i \kappa_{j, n}\left(-\alpha_{j, n}+\beta_{j, n}\right) B_{j, n} \Phi_{j, n}(y, z)=-i \rho_{0} \omega \hat{V}_{s}(y) \\
\sum_{n=0}^{\infty} \sum_{j=0}^{\infty}\left[\alpha_{j, n}\left(1-\zeta_{L} \frac{\kappa_{j, n}}{k}\right) e^{-i \kappa_{j, n} L}+\beta_{j, n}\left(1+\zeta_{L} \frac{\kappa_{j, n}}{k}\right) e^{+i \kappa_{j, n} L}\right] B_{j, n} \Phi_{j, n}(y, z)=0 \tag{10.2.5}
\end{gather*}
$$

It might appear that the preceding constitutes two equations for many unknown $\alpha_{j, n}$ and $\beta_{j, n}$ values, but such thinking ignores the orthogonality of the transverse mode functions. Each equation is multiplied by an arbitrarily selected $\Phi_{m, \ell} \equiv B_{m, \ell} \phi_{y, m} \phi_{z, \ell}$ and integrated over the cross section, $0<y<H, 0<z<W$. These operations filter out all term in the summations for which $n \neq m$ and $j \neq \ell$ remain. What remains is a pair of equations,

$$
\begin{gather*}
\alpha_{m, \ell}-\beta_{m, \ell}=\rho_{0} c \frac{k}{\kappa_{m, \ell}} V_{m, \ell}  \tag{10.2.6}\\
\alpha_{m, \ell}\left(1-\zeta_{L} \frac{\kappa_{m, \ell}}{k}\right) e^{-i \kappa_{m, \ell} L}+\beta_{m, \ell}\left(1+\zeta_{L} \frac{\kappa_{m, \ell}}{k}\right) e^{+i \kappa_{m, \ell} L}=0
\end{gather*}
$$

The $V_{m, \ell}$ coefficients are those of a modal series for $V_{s}$ defined in Eq.(9.3.37), specifically

$$
\begin{equation*}
V_{m, \ell}=\frac{1}{H W} \int_{0}^{W} \int_{0}^{H} \Phi_{m, \ell}(y, z) V_{s}(y, z) d y d z \tag{10.2.7}
\end{equation*}
$$

Orthogonality has reduced the problem to that of solving two simultaneous equations for $\alpha_{m, \ell}$ and $\beta_{m, \ell}$, which leads to

$$
\begin{align*}
\alpha_{m, \ell} & =\frac{\left(1+\zeta_{L} \frac{\kappa_{m, \ell}}{k}\right) e^{+i \kappa_{m, \ell} L}}{\Delta(k, j, n)} \rho_{0} c \frac{k}{\kappa_{m, \ell}} V_{m, \ell} \\
\beta_{m, \ell} & =\frac{\left(-1+\zeta_{L} \frac{\kappa_{m, \ell}}{k}\right) e^{-i \kappa_{m, \ell} L}}{\Delta(k, j, n)} \rho_{0} c \frac{k}{\kappa_{m, \ell}} V_{m, \ell} \tag{10.2.8}
\end{align*}
$$

where $\Delta(k, j, n)$ is the determinant of Eq. (10.2.6),

$$
\begin{equation*}
\Delta\left(\kappa_{j, n}\right)=\left(1+\zeta_{L} \frac{\kappa_{j, n}}{k}\right) e^{i \kappa_{j, n} L}+\left(1-\zeta_{L} \frac{\kappa_{j, n}}{k}\right) e^{-i \kappa_{j, n} L} \tag{10.2.9}
\end{equation*}
$$

Substitution of the solution into the modal series, Eq. (10.2.2), leads to

$$
\begin{align*}
P & =\rho_{0} c \sum_{n=0}^{\infty} \sum_{j=0}^{\infty} \frac{k V_{j, n}}{\kappa_{j, n} \Delta\left(\kappa_{j, n}\right)}\left[\left(1+\zeta_{L} \frac{\kappa_{j, n}}{k}\right) e^{i \kappa_{j, n}(L-x)}\right.  \tag{10.2.10}\\
& \left.-\left(1-\zeta_{L} \frac{\kappa_{j, n}}{k}\right) e^{-i \kappa_{j, n}(L-x)}\right] \Phi_{j, n}(y, z)
\end{align*}
$$

The fact that the variable for the axial dependence in Eq. (10.2.10) is $L-x$ highlights the reflection process at $x=L$. Another manifestation of that attribute comes from factoring $1+\zeta_{L} \kappa_{j, n} / k$ out of the bracketed term. Doing so displays the role of a modal reflection coefficient,

$$
\begin{equation*}
\left(R_{L}\right)_{j, n}=\frac{\zeta_{L} \frac{\kappa_{j, n}}{k}-1}{\zeta_{L} \frac{\kappa_{j, n}}{k}+1} \tag{10.2.11}
\end{equation*}
$$

so that

$$
\begin{equation*}
P=\rho_{0} c \sum_{n=0}^{\infty} \sum_{j=0}^{\infty} V_{j, n} \frac{k}{\kappa_{j, n}}\left[\frac{e^{i \kappa_{j, n}(L-x)}+\left(R_{L}\right)_{j, n} e^{-i \kappa_{j, n}(L-x)}}{e^{i \kappa_{j, n} L}-\left(R_{L}\right)_{j, n} e^{-i \kappa_{j, n} L}}\right] \Phi_{j, n}(y, z) \tag{10.2.12}
\end{equation*}
$$

Setting $\Phi_{j, n}(y, z)=1$ and $\kappa_{j, n}=k$ would make each term in the summation be like Eq. (3.2.21) for plane waves.

Some individuals prefer a form of Eqs. (10.2.9) and (10.2.10) that results when Euler's identity is used to remove the complex exponentials in favor of sine and cosine functions. The resulting expressions are not displayed here because such forms offer no merit if $\zeta_{L}$ or $\kappa_{j, n}$ is complex.

Resonances correspond to values of $k$ for which $\Delta\left(\kappa_{j, n}\right)=0$. It can proven that this condition occurs only if the impedance $\zeta_{L}$ at $x=L$ and the impedances of the sidewalls are infinite, zero, or purely imaginary, that is, purely reactive. As was shown earlier, any frequency at which $\Delta=0$ is a natural frequency of the cavity. If any impedance has a nonzero real part, energy is absorbed by that wall, thereby inhibiting unlimited growth of the pressure amplitude. However, if the resistive part of each impedance is substantially smaller than $\rho_{0} c$, a plot of pressure as a function of frequency will exhibit prominent peaks in the vicinity of the natural frequencies associated with zero resistive parts.

When any wall impedance is a finite compliance or inertance, determination of the natural frequencies by solving the eigenvalue problem is not a simple task. The difficult lies in the fact that a transverse eigenvalue, $\mu_{j}$ or $\nu_{n}$, will depend on the frequency because the associated characteristic equation contains $k$, see, for example,

Eq. (9.2.32). Concurrently, the axial wavenumber $\kappa_{j, n}$ must satisfy its own characteristic equation, specifically Eq.(10.2.8). In addition, the natural frequency is related to the eigenvalues through Eq. (10.2.3), so that

$$
\begin{equation*}
\left(k_{\mathrm{nat}}\right)_{j, n}=\left(\kappa_{j, n}^{2}+\mu_{j}^{2}+\nu_{n}^{2}\right)^{1 / 2} \tag{10.2.13}
\end{equation*}
$$

Thus, solution of the eigenvalue problem in this situation requires simultaneous solution of three characteristic equations and one algebraic equation. The unknowns are the three wavenumbers and the $k_{\text {nat }}$. Such a solution could be obtained by implementing a four variable Newton's method. A simpler, though less efficient, procedure is to sweep through a range of $k$. For each value, each characteristic equation is solved for the associated wavenumber. If these values are consistent with Eq.(10.2.13), then that value of $k$ marks a natural frequency. Regardless of the procedure that is implemented, care must be exercised to obtain all natural frequencies in the interval of interest.

One of the useful aspects of a series representation of the pressure field is the uncoupled nature of the equations to be solved. That is, we can find the amplitude coefficients $\alpha_{j}$ and $\beta_{j}$ for a mode without finding the coefficients of the other modes. In some situations, it might be that more than one wall moves. In that case, the response may be found by superposing solutions. In each subproblem, one wall moves and the other vibrating walls are held fixed; that is, they become rigid. The axial direction for each subproblem would be normal to the wall that is vibrating. Addition of the pressure from each subproblem yields the pressure field.

> EXAMPLE 10.1 All walls of a rectangular cavity are rigid. The dimensions are $L$ in the $x$ direction, $H$ in the $y$ direction, and $W$ in the $z$ direction. The upper half of the wall at $x=0$ is stationary. The lower half is a rigid plate having mass $M$ that is forced to undergo an oscillation at frequency $\omega$. Thus, the end condition may be written as $v_{x}=\operatorname{Re}\left(V_{0} \exp (-i \omega t)\right)[h(y)-h(y H / 2)]$, where $h$ denotes the step function. Derive an expression for the amplitude of the force that must be applied to the plate. Evaluate this force as a function of the nondimensional frequency $k L$, for the case where $H / L=0.2$ and $W / L=0.1$. The mass of the plate equals the mass of the fluid contained in the cavity.

## Significance

This basic operations for the analysis are fairly straightforward. What makes the problem interesting is the interpretation of the nature of the forces acting on the plate, which will afford a different perspective to the meaning of resonances.

## Solution

The force acting on the plate must overcome the inertia of the plate and the resultant force exerted by the pressure. This resultant, which acts in the negative $x$ direction if the pressure is positive, is the integral over the plate of the pressure at $x=0$. Thus, the complex amplitude of the pressure resultant is

$$
\begin{equation*}
F_{\mathrm{ac}}=\left.\int_{0}^{W} \int_{0}^{H / 2} P\right|_{x=0} d y d z \tag{1}
\end{equation*}
$$

The acceleration of the plate is $\operatorname{Re}\left(i \omega V_{0} \exp (i \omega t)\right)$. Thus, the force amplitude required to translate the plate is

$$
\begin{equation*}
F=i \omega M V_{0}+F_{\text {ac }} \tag{2}
\end{equation*}
$$

Now, we turn our attention to the fluid. The walls that coincide with planes of constant $y$ and $z$ are not forced, so we take them to be the sidewalls of the waveguide. The corresponding transverse mode functions are

$$
\begin{equation*}
\Phi_{j, n}=B_{j, n} \cos \left(\frac{j \pi y}{H}\right) \cos \left(\frac{n \pi z}{W}\right), \quad j, n=0,1,2, \ldots \tag{3}
\end{equation*}
$$

Normalization of these functions leads to

$$
B_{j, n}=\sqrt{\frac{4}{\left(1+\delta_{j, 0}\right)\left(1+\delta_{n, 0}\right)}}
$$

The axial wavenumber is

$$
\begin{equation*}
\kappa_{j, n}=\left[k^{2}-\left(\frac{j \pi}{H}\right)^{2}-\left(\frac{n \pi}{W}\right)^{2}\right]^{1 / 2} \tag{4}
\end{equation*}
$$

with the usual proviso that $\kappa_{j, n}$ is negative imaginary if the quantity within the brackets is negative.

In Eq. (10.2.4), the axial velocity imposed at $x=0$ is $V_{s}(y, z)$. Here, this function is specified to be

$$
V_{s}=V_{0}[h(y)-h(y-H / 2)], \quad 0<y<H, 0<z<W
$$

The corresponding velocity coefficients obtained from Eq. (10.2.7) are

$$
\begin{align*}
V_{j, n} & =V_{0} B_{j, n} \mathcal{I}_{j, n} \\
\mathcal{I}_{j, n} & =\frac{1}{H W} \int_{0}^{W} \int_{0}^{H / 2} \cos \left(\frac{j \pi y}{H}\right) \cos \left(\frac{n \pi z}{W}\right) V_{0} d y d z \\
& =\left\{\begin{array}{l}
0 \text { if } n>0 \\
\frac{1}{2} \text { if } j=n=0 \\
\frac{1}{j \pi} \sin \left(\frac{j \pi}{2}\right) \text { if } j>0 \text { and } n=0
\end{array}\right. \tag{5}
\end{align*}
$$

The $V_{j, n}$ coefficients are zero if $n>0$ because the imposed velocity is constant in the $z$ direction, and a constant is orthogonal to the $z$ dependence of all transverse
mode functions for $n>0$. The velocity distribution is the excitation, so any velocity coefficient that is zero means that the corresponding mode is not excited. In addition, the factor $\sin (j \pi / 2)$ is zero if $j$ is even, so $V_{j, n} \equiv 0$ for even $j$. Thus, rather than requiring evaluation of a double sum, Eq. (10.2.10) reduces to a single sum over $j=0,1,3, \ldots$ with $n=0$. The only other adaptation required to use that equation is to let $\zeta_{L}$ be infinite. The resulting series for the pressure field is

$$
\begin{equation*}
P=\rho_{0} c V_{0} \sum_{j=0,1,3, \ldots}^{\infty}\left(B_{j, 0}\right)^{2} \mathcal{I}_{j, 0}\left(\frac{k L}{\kappa_{j, 0} L}\right)\left(\frac{e^{i \kappa_{j, n}(L-x)}+e^{-i \kappa_{j, n}(L-x)}}{e^{i \kappa_{j, n} L}-e^{-i \kappa_{j, n} L}}\right) \cos \left(\frac{j \pi y}{H}\right) \tag{6}
\end{equation*}
$$

The last step is to form the resultant force by substitution of this expression, evaluated at $x=0$, into Eq.(1). The integral for this evaluation is the same as the one for $\mathcal{I}_{m, \ell}$ in Eq. (5). When the complex exponential sum and difference are replaced by sinusoidal functions, the pressure resultant reduces to

$$
\begin{gather*}
F_{\mathrm{ac}}=H W Z_{\mathrm{ac}} V_{0} \\
Z_{\mathrm{ac}}=i \rho_{0} c \sum_{j=0,1,3, \ldots}^{\infty}\left(B_{j, 0}\right)^{2}\left(\mathcal{I}_{j, 0}\right)^{2}\left(\frac{k L}{\kappa_{j, 0} L}\right) \cot \left(\kappa_{j, 0} L\right) \tag{7}
\end{gather*}
$$

An expression for the force that must be applied to the plate is obtained by substituting $F_{\text {ac }}$ into Eq. (2), which results in

$$
\begin{equation*}
F=H W Z_{F} V_{0} \tag{8}
\end{equation*}
$$

The factor $Z_{F}$ is the total impedance seen by $F$. It consists of a contributions from the acoustic pressure and the inertia of the plate,

$$
\begin{equation*}
Z_{F}=Z_{\mathrm{ac}}+Z_{M}, \quad Z_{M}=\frac{i \omega M}{H W} \equiv i \rho_{0} c k L\left(\frac{M}{\rho_{0} L W H}\right) \tag{9}
\end{equation*}
$$

Because of the simplicity of this system, the value of $Z_{F} /\left(\rho_{0} c\right)$ depends on only three parameters: the nondimensional frequency $k L$, the ratio of the plate's mass to that of the fluid contained within the cavity $M /\left(\rho_{0} L W H\right)$, and $H / L$. If $j$ is less than the cutoff value $k H / \pi$, then $\kappa_{j, 0}$ is real, whereas $\kappa_{j, 0}$ is negative imaginary if $j$ exceeds the cutoff value. In either case, the terms inside the summation are real. It follows that both $Z_{\text {ac }}$ and $Z_{M}$ are reactive at any frequency. The corollary is that the time-averaged power input to the cavity by $F$ is zero. This is as it must be, because the walls are rigid and therefore cannot dissipate energy.

The value of $j$ at which the series in Eq. (7) may be truncated is readily identified from Eq. (4), which indicates that $\kappa_{j, 0} L \approx-i j \pi(H / L)$ when $j \pi / H \gg k$. For very large $j$, the terms in the summation are proportional to $\operatorname{coth}(j \pi(H / L))$, which approaches unity. The coefficients $B_{j, 0}=2^{1 / 2}$ and $\mathcal{I}_{j, 0}$ are inversely proportional to $j \pi$. It follows that for large $j$ the terms in the summation are inversely proportional
to $j^{3}$. Thus, a truncation criterion that $j>k L(H / L) / \pi$ and $j>20$ should be adequate. Figure 1 shows the result of computing $Z_{F}$ for the given parameters with the series truncated at $j=21$. In order to understand the role of the resultant acoustic force, the frequency dependence of $Z_{\text {ac }}$ also is shown in Fig. 1. The second graph gives a finer picture of the behavior in the vicinity of the first resonances.


Figure 1.

Resonances occur if the denominator in Eq. (6) vanishes for any $j$. The denominator may be written as $2 i \sin \left(\kappa_{j, 0} L\right)$. Equating it to zero yields $\kappa_{j, 0} L=m \pi$. Then, Eq. (4) yields the natural frequencies,

$$
\begin{equation*}
(k L)_{\mathrm{nat}}=\left[(m \pi)^{2}+(j \pi)^{2}\left(\frac{H}{L}\right)^{2}\right]^{1 / 2}, m=0,1,2, \ldots, \quad j=0,1,3, \ldots \tag{10}
\end{equation*}
$$

These values are marked with an asterisk in Fig. 1. The values in ascending order are tabulated below.

| $j$ | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 0 |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $m$ | 0 | 1 | 2 | 3 | 4 | 5 | 0 | 1 | 2 | 3 | 6 |
| $(k L)_{\text {nat }}$ | 0.00 | 3.14 | 6.28 | 9.42 | 12.57 | 15.71 | 15.71 | 16.02 | 16.92 | 18.32 | 18.85 |

Two different $j, m$ pairs give rise to the same natural frequency at $k L=15.71$, and the spacing between natural frequencies decreases with increasing frequency. Both of these features will be explored in detail when we analyze natural cavity modes.

As expected, $\left|Z_{F}\right|$ is very large in the vicinity of a natural frequency, whereas the value of $Z_{M}$ is an inertance that is proportional to the frequency. The plots in Fig. 1 indicate that $Z_{M}$ is the dominant effect away from a resonance. As the frequency is increased beyond a resonance, $Z_{\text {ac }}$ transitions to negative imaginary; that is, it
is a compliance. Further increase toward resonance results in a sharp increase of this compliance, eventually becoming singular as a resonance is approached from below. The consequence is that at some frequency between resonances the value of $Z_{F} \equiv Z_{M}+Z_{\text {ac }}$ is zero. This has an important implication.

Let us change our perspective to ask what the plate's displacement would be if we were to drive it with a specified $F$. The displacement is $U=V_{0} /(i \omega)$, so Eq. (8) tells us that

$$
\begin{equation*}
U=\frac{F}{i \omega H W}\left(\frac{1}{Z_{F}}\right) \tag{11}
\end{equation*}
$$

From this viewpoint, a large value of $Z_{F}$ at a resonance leads to a very small displacement. In contrast, a very small value of $Z_{F}$ leads to a large displacement. This is a resonance of the coupled system formed by the plate and the cavity. We say that it is a fluid-structure resonance.

This might seem to be somewhat confusing, but the two situations are readily distinguished if we think of an experiment. The situation implied in the problem statement is that a machine exerts a force amplitude $F$ that can be as large as necessary to maintain a specified value of $V_{0}$ regardless of the pressure within the cavity. In this scenario, the pressure will grow greatly when the frequency is close to any $(k L)_{\text {nat }}$, so the value of $F$ required to sustain $V_{0}$ will also be very large. In another experiment, the machine applies the constant amplitude force $F$. In that case, Eq. (11) gives the displacement $U$. Setting $V_{0}=i \omega U$ in Eq. (7) gives the corresponding acoustic force $F_{\text {ac }}$ applied to the plate,

$$
\begin{equation*}
F_{\mathrm{ac}}=\frac{Z_{\mathrm{ac}}}{Z_{M}+Z_{\mathrm{ac}}} F \tag{12}
\end{equation*}
$$

The value of $F_{\text {ac }} /(H W)$ serves as a metric for the pressure within the cavity. In the second experiment, the displacement and pressure within the cavity are very large at frequencies for which $Z_{M}+Z_{\mathrm{ac}}=0$; this is the fluid-structure resonance condition. The contrasting situation places the frequency at one of the natural frequencies of the cavity, as given in Eq. (10). At these frequencies, $Z_{\mathrm{ac}}$ is very large. Equation (11) gives a small value for $U$ in this case, and Eq. (12) gives $F_{\text {ac }}=F$. It follows that when we talk about a resonance, we must be unambiguous in specifying how the system is driven. To some extent, we encountered this same issue when we considered closure of one-dimensional waveguides, specifically in Example 3.4.

EXAMPLE 10.2 A rectangular tank of water has width and height $4 a$. The bottom and side walls are rigid, and the top is a free surface. A circular piston transducer whose radius is $a$ is flush mounted on the rigid wall at $x=L$, so that $V_{x}=V_{0} \exp (i \omega t)$ on the piston face and zero otherwise. Derive an expression for the pressure at an arbitrary location as a function of $k a$. Then evaluate $|p| /\left(\rho_{0} c V_{0}\right)$ at $x=0, L / 2$, and $L$ on the centerline, $y=z=2 a$, for the frequency range $0<k a<8$. Also evaluate the face-averaged impedance seen by the transducer. The aspect ratio for the evaluation is $L / a=5$.


Figure 1.

## Significance

From a conceptual viewpoint, the system in this example is very much like the one analyzed in the previous example. The primary difference is the fact that forced cavity modes covering a range of wavenumbers in both the $y$ and $z$ directions are excited here. The other significant difference from the previous example is that numerical methods will be required to evaluate the modal velocity coefficients. A different view of the behavior at resonances will emerge.

## Solution

Identification of symmetry properties is a useful starting point. The vertical plane that contains the centerline cuts the tank in half, and the walls have the same properties on either side of this plane. Furthermore, the input velocity field is symmetric with respect to this plane. Consequently, the pressure field must share this symmetry. The normal to the symmetry plane is the $y$ direction, so only forced cavity modes that are even functions with respect to $y=2 a$ will participate in the response. The walls at $y=0$ and $y=4 a$ are rigid. The full set of two-dimensional modes for rigid walls separated by distance $H$ is $\cos (s \pi / H), s=0,1,2, \ldots$ The symmetric ones correspond to even $s$. In contrast, there is no symmetry with respect to any horizontal plane, so the full set of rigid/pressure-release two-dimensional modes are required to describe the field's dependence on $z$. The length of both sides of the cross section is $4 a$, so the normalized transverse modes for the analysis are

$$
\begin{gather*}
\Phi_{j, n}=B_{n, j} \cos \left(\frac{2 j \pi y}{4 a}\right) \cos \left(\frac{(2 n-1) \pi z}{2(4 a)}\right) ; \quad j=0,1,2, \ldots, n=1,2, \ldots \\
B_{j, n}=\sqrt{\frac{4}{1+\delta_{j n, 0}}} \tag{1}
\end{gather*}
$$

The associated axial wavenumbers are

$$
\begin{equation*}
\kappa_{j, n}=\left[k^{2}-\left(\frac{j \pi}{2 a}\right)^{2}-\left(\frac{(2 n-1) \pi}{8 a}\right)^{2}\right]^{1 / 2} \tag{2}
\end{equation*}
$$

As always, the square root should be such that $\kappa_{j, n}$ is either positive real or negative imaginary.

For the sake of variety, the boundary excitation has been situated at $x=L$. The piston's axial velocity is $V_{x}=V_{0}$ over its face and $V_{x}=0$ for points that are not on the piston. Consequently, the integrand for the modal velocity coefficients is zero outside the piston, so that

$$
\begin{equation*}
V_{j, n}=\frac{1}{16 a^{2}} \int_{0}^{4 a} \int_{0}^{4 a} V_{x} \Phi_{j, n} d y d z=\frac{V_{0}}{16 a^{2}} \iint_{\mathcal{A}_{\text {piston }}} \Phi_{j, n} d y d z \tag{3}
\end{equation*}
$$

In view of the circular shape of the piston's face, it is logical to use polar coordinates centered on the piston to formulate the integral. The coordinate transformation is

$$
y=2 a+R \cos \theta, z=2 a+R \sin \theta
$$

The face of the piston corresponds to $0<R<a,-\pi<\theta<\pi$, which converts the velocity coefficients to

$$
\begin{align*}
V_{j, n}= & \frac{1}{16 a^{2}} \int_{0}^{a} \int_{-\pi}^{\pi} B_{n, j} \cos \left(\frac{j \pi}{2}\left(2+\frac{R}{a} \cos \theta\right)\right)  \tag{4}\\
& \times \cos \left(\frac{\pi}{8}(2 n-1)\left(2+\frac{R}{a} \sin \theta\right)\right) R d \theta d R
\end{align*}
$$

No analytical result for these integrals is readily available, so they are evaluated numerically. Some software packages contain a double integration routine. (In MATLAB, it is dblquad.) Otherwise, the integration may be performed by dividing a circle into a set of small patches subtending $\Delta R^{\prime}$ and $\Delta \theta$. The double integral is approximately the sum of the value of $\Phi_{n, j}$ at the center point of each patch, multiplied by the area of the patch. Figure 2 plots the values $V_{n, 1}$ and $V_{0, j}$. It can be seen that the coefficients decrease, but slowly and in an oscillatory manner.


Figure 2.

The values of $V_{j, n}$ are the last set of parameters required to form the modal series, Eq. (10.2.10). If we wish to employ the derived formulas, we must recognize that the $x$ direction used for the derivation is reversed from the one we have used for the
present system, with the excitation occurring at $x=L$ rather than $x=0$. We may adjust for this difference by replacing $x$ in the formulas with $L-x$. The undriven end is rigid, so we let $\zeta_{L}$ be infinite in Eq.(10.2.10). This is handled by dropping terms that do not contain $\zeta_{L}$ in the numerator and denominator, which is $\Delta\left(\kappa_{j, n}\right)$ in Eq. (10.2.9). Thus, the modal pressure series is transformed into

$$
\begin{align*}
P & =\sum_{j=0}^{\infty} \sum_{n=1}^{\infty} \frac{k}{\kappa_{j, n}} \rho_{0} c V_{n, j}\left(\frac{e^{-i \kappa_{j, n} x}+e^{+i \kappa_{j, n} x}}{e^{+i \kappa_{j, n} L}-e^{-i \kappa_{j, n} L}}\right) \Phi_{n, j}(y, z)  \tag{5}\\
& \equiv \sum_{j=0}^{\infty} \sum_{n=1}^{\infty} \frac{k}{\kappa_{j, n}} \rho_{0} c V_{n, j} \frac{\cos \left(\kappa_{j, n} x\right)}{i \sin \left(\kappa_{j, n} L\right)} \Phi_{n, j}(y, z)
\end{align*}
$$

Resonances correspond to the zeros of the denominator. This is examined in detail in the discussion.

The face-averaged impedance, which is the average pressure on the piston face divided by the average velocity, may be computed directly from the other parameters. The average velocity amplitude on the piston is $V_{0}$. The average pressure is found as an integral over the face of the piston,

$$
P_{\mathrm{av}}=\left.\frac{1}{\pi a^{2}} \iint_{\mathcal{A}_{\text {piston }}} P\right|_{x=L} d y d z
$$

The pressure is described by the modal series, Eq. (5), which we evaluate at $x=L$. In order to better understand the magnitude of the terms, we shall replace the complex exponentials with equivalent forms derived from Euler's identity. Thus, we form

$$
\begin{equation*}
P_{\mathrm{av}}=\frac{1}{\pi a^{2}} \sum_{j=0}^{\infty} \sum_{n=1}^{\infty} \frac{k a}{i \kappa_{n, j} a} \rho_{0} c V_{j, n} \cot \left(\kappa_{j, n} L\right) \iint_{\mathcal{A}_{\text {piston }}} \Phi_{j, n}(y, z) d y d z \tag{6}
\end{equation*}
$$

The integral appeared previously in Eq. (3). It follows that

$$
\iint_{\mathcal{A}_{\text {piston }}} \Phi_{j, n}(y, z) d y d z=16 a^{2} \frac{V_{j, n}}{V_{0}}
$$

The ratio of pressure to velocity is an impedance. Thus, the average acoustic impedance seen by the piston may be determined by evaluating

$$
\begin{equation*}
Z_{\mathrm{ac}}=\frac{P_{\mathrm{av}}}{V_{0}}=16 \sum_{j=0}^{\infty} \sum_{n=1}^{\infty} \frac{k a}{i \kappa_{j, n} a} \rho_{0} c\left(\frac{V_{j, n}}{V_{0}}\right)^{2} \cot \left(\kappa_{n, j} L\right) \tag{7}
\end{equation*}
$$

Evaluation of the pressure and average impedance requires computation of double sums over the $n$ and $j$ mode numbers. Let $J=\max (j)$ and $N=\max (n)$ be the numbers at which the summations are halted. One criterion is that $J$ and $N$ may be
set such that $V_{j, n}$ is negligible for $j>J$ and $n>N$. The slow oscillatory reduction of $V_{j, n}$ indicated in Fig. 2 suggests that this criterion would require that $J$ and $N$ be quite large. Another possibility is that the series lengths may be set according to the behavior of $\kappa_{j, n} a$. Examination of Eqs. (5) and (7) shows that convergence of the pressure and impedance will be attained if the $\kappa_{n, j}$ values for omitted modes are large in magnitude. This condition will be attained if $J$ is sufficiently large that $\kappa_{J, n}$ is negative imaginary regardless of $n$, and $N$ is sufficiently large that $\kappa_{j, N} a$ is negative imaginary regardless of $j$. Equation (2) for $\kappa_{j, n}$ shows that this condition can be met if $J>2 k a / \pi$ and $N>4 k a / \pi+1 / 2$. The highest frequency for the computations is $k a=8$, which gives $J>6$ and $N>11$. Doubling these values is more than adequate because $\kappa_{J, 1} a \approx-17 i$ and $\kappa_{0, N} a \approx-15 i$, whereas the smallest values of $\left|\kappa_{n, j}\right|$ are close to zero for modes that are closest to the cutoff condition. A by-product of selecting $J$ and $N$ in this manner is certainty that all modes that might be resonant are included. This is so because $\kappa_{n, j}$ of all omitted modes are negative imaginary, so the exponential terms in the denominators are real.

The computational algorithm developed for a rectangular waveguide in Example 9.7 is equally applicable here. Preliminary to the main frequency loop, all quantities that are independent of frequency are evaluated. This entails setting up an outer loop over $j=0, \ldots, J$ and an inner loop over $n=1, \ldots, N$. The element index is set at $m=j N+n$. Within these loops, the values of $\tilde{B}_{m} \equiv B_{j, n}$ are computed according to Eq. (1), and the double integral for $\tilde{V}_{m} \equiv V_{j, n} / V_{0}$ is carried out. We are working nondimensionally, so the latter coefficients are given by

$$
\tilde{V}_{m}=\frac{1}{16} \int_{0}^{1} \int_{-\pi}^{\pi} \Phi_{j, n}(2+\cos \theta, 2+\sin \theta) R^{\prime} d \theta d R^{\prime}
$$

Another item that may be computed independently of the frequency is the value of each transverse mode function on the centerline, $y=z=2 a$, at which the pressure will be computed. These values form $\tilde{\Phi}_{m} \equiv \Phi_{j, n}(2 a, 2 a)$, which are stored in a diagonal matrix [ $\Phi_{m}$ ].

The frequency loop finely increments the $k a$ value between 0 and 8 . At each frequency, the loops over $j=0, \ldots, J$ and $n=1, \ldots, N$ are repeated. The index $m=j N+n$ is set as the subscript for storing arrays as column vectors. The nondimensional axial wavenumber $\tilde{\kappa}_{m} a$ is obtained from Eq. (2). The pressure at the designated points $x_{1}=0, x_{2}=L / 2 \sim$, and $x_{3}=L$ and the impedance value are computed by placing the pressure in Eq. (5) and the impedance coefficient in Eq. (7) into matrix form as

$$
\begin{align*}
\left\{\frac{P}{\rho_{0} c V_{0}}\right\} & =[E][D]_{\text {diag }}\left[\tilde{\Phi}_{m}\right]_{\text {diag }}\{\tilde{V}\}  \tag{8}\\
\frac{Z_{\mathrm{av}}}{\rho_{0} c} & =\{\tilde{V}\}^{\mathrm{T}}[F]_{\text {diag }}\{\tilde{V}\}
\end{align*}
$$

where

$$
\begin{gathered}
D_{m, m}=\frac{k a}{i \kappa_{m} a}\left(\frac{1}{\sin \left(\left(\tilde{\kappa}_{m} a\right) \frac{L}{a}\right)}\right) \\
E_{s, m}=\cos \left(\left(\tilde{\kappa}_{m} a\right) \frac{L}{a} \frac{x_{s}}{L}\right) \\
F_{m, m}=\frac{k a}{i \kappa_{m} a} \cot \left(\left(\tilde{\kappa}_{m} a\right)\left(\frac{L}{a}\right)\right)
\end{gathered}
$$

The pressure amplitude at the designated locations is plotted as a function of frequency in Fig.3. A casual viewing might lead one to conclude that each plot describes random noise. That this is not so is evidenced by the behavior in the lower frequency range, where distinct resonance peaks can be identified. Each of the highfrequency peaks also is caused by a resonance. The denominators in Eqs. (5) and (7) may be written equivalently as $2 i \sin \left(\kappa_{j, n} a(L / a)\right)$. Resonance corresponds to this quantity being zero, which occurs if $\kappa_{j, n} a=(a / L) s \pi, s=0,1,2, \ldots$. In view of the definition of $\kappa_{j, n}$ in Eq. (2), the resonance condition is marked by

$$
\begin{equation*}
\left(\frac{n \pi}{2}\right)^{2}+\left(\frac{(2 j-1) \pi}{8}\right)^{2}+\left(\frac{a}{L} \pi\right)^{2} s^{2}=(k a)^{2} \tag{9}
\end{equation*}
$$

As the frequency increases, it becomes increasingly likely that there will be one or more $n, j$ pairs that satisfy Eq. (9). For example, the first resonance occurs at $k a=\pi / 8=0.393$ corresponding to $n=0, j=1$, and $s=0$. This is contrasted by the situation for $k a=7.54$, for which $\left|\sin \left(\kappa_{j, n} a(L / a)\right)\right|<0.05$ for $n=2, j=4$ $(s=10), n=1, j=9(s=5)$, and $n=j=4(s=5)$. The fact that the resonances become very close is further evidenced by the fact that a slight shift to $k a=7.55$ leads to $\left|\sin \left(\kappa_{j, n} a(L / a)\right)\right|<0.05$ for $n=0, j=1(s=12)$, as well as $n=2$, $j=4(s=10)$.


Figure 3.
Figure 4 zooms in on the low-frequency range. With one exception, a resonance frequency is manifested by the pressure at each location showing a peak. The exception occurs at $k a=0.74$, where the pressure at $x=L / 2$ does not show a peak. This is a consequence of the fact that the resonant $(0,1)$ mode corresponds to
$\cos \left(\tilde{\kappa}_{m}(x / L)(L / a)\right)=0$ at $x=L / 2$. Hence, its does not contribute to the pressure at this location even though the amplitude of this mode is large. Unlike resonances, sharp minima of $|P|$ are not correlated between different locations. These occur because the many terms in the modal sums accumulate to a very small value, and the terms in that sum depend on position.


The $V_{n, j}$ coefficients given in Eq. (4) are real, so $Z_{\text {ac }}$ is imaginary. This means that $Z_{\mathrm{ac}}$ is reactive, as it was in the previous example. The average impedance on the piston face is depicted in Fig. 5. (The dashed lines describe the discontinuous drop from an infinite inertance to zero compliance as a natural frequency is passed.) The overall appearance of Fig. 5 is like Fig. 1 of the previous example, except that the spacing between natural frequencies is irregular here.


Figure 5.

### 10.2.2 Spherical Cavities

One seldom encounters a spherical room, although the author does recall visiting a combat flight simulator at an aerospace company in which images were projected on a nearly complete spherical wall. Nevertheless, being able to analyze waves in an enclosed spherical space affords us the opportunity to examine the differences between different types of cavities and thereby understand how curved boundaries affect basic phenomena. We first encountered a spherical cavity in Chap. 6. The earlier treatment was limited to radially symmetric waves. The development of spherical harmonics in Chap. 7 makes it possible to generalize the analysis.

It might seem that a spherical domain is unlike the interior of a box, but there are many similarities. Although a sphere has one physical boundary, whereas a box has six walls, there are additional boundary conditions for the former that play the roles
of walls. One we encountered in the study of radially symmetric waves, wherein a primary consideration is that the general solution of the wave equation must not be singular at the center. Other conditions are continuity. If we know the complex amplitude $P$ at a set of spherical coordinates $(r, \psi, \theta)$, an increase or decrease of $\theta$ by $2 \pi$ must yield the same values for the pressure and particle velocity. (This condition is identically met by an axisymmetric field, which is the situation we will consider.) Continuity also is an issue at the poles, because the pressure and particle velocity at $\psi=0$ and $\psi=\pi$ must be finite.

Fortunately, these issues are addressed by a spherical harmonic series representation. Thus, the starting point for our exploration is the general solution for axisymmetric waves in a spherical geometry. This was the subject of Sect.7.1.

## Axisymmetric Excitation

We wish to determine the pressure in the region enclosed by a vibrating spherical wall. The wall's motion is harmonic. A further restriction is that the motion is limited to being axisymmetric. In accord with our standard practice we designate the symmetry axis as $z$. Axisymmetry of the excitation leads to axisymmetry of the response, so the pressure field is $p(\bar{x}, t)=\operatorname{Re}[P(r, \psi) \exp (i \omega t)]$. The general solution of the Helmholtz equation in axisymmetric spherical coordinates is a series of spherical harmonics. Each harmonic is a product of a radial function and a polar function.

The polar angle is $\psi$, but the mathematical variable that was used in Chap. 7 was the transformed variable $\eta=\cos \psi$. Some individuals view $\eta$ as the projection onto the $z$ axis of a unit radial line at angle $\psi$. The angular functions are Legendre polynomials, $P_{m}(\eta)$. When $\eta$ is replaced by its definition, the result is that the angular functions are $P_{m}(\cos \psi)$. These are said to be Legendre functions. An analysis of the Helmholtz equation usually will proceed equally well using either $\psi$ or $\eta$.

Radiation exterior to a sphere causes waves to propagate outward. A spherical Hankel function $h_{m}^{(2)}(k r)$ is appropriate to this situation because at long distances it approaches a proportionality to $\exp (-i k r)$. This function is not appropriate by itself for the interior problem because it is singular at the origin, $r=0$. The singularity of this solution could be canceled by adding the Hankel function $h_{m}^{(1)}(k r)$ for an inward propagating spherical wave. Such a representation would lead to an analysis that parallels the analysis of a rectangular cavity. Finiteness at $r=0$ would be the analog of the boundary condition for the wall opposite the driven one in the rectangular geometry.

The merits of seeing this analogy are counterbalanced by the complications associated with dealing with (complex) Hankel functions. The Helmholtz equation is linear and real, so the real and imaginary parts of each Hankel function satisfy that equation individually. The singularity at the origin occurs in the imaginary part, so we discard it. The result is that the radial part of a spherical harmonic for an enclosed region is a spherical Bessel function $j_{m}(k r)$. The order $m$ is the same as the degree of one of the Legendre functions.

A general solution for the field within a sphere is obtained by multiplying the product of each Legendre function and the corresponding spherical Bessel function by a coefficient whose value must be determined. Thus, the basic ansatz for the complex pressure amplitude is

$$
\begin{equation*}
P=\sum_{m=0}^{\infty} C_{m} j_{m}(k r) P_{m}(\cos \psi) \tag{10.2.14}
\end{equation*}
$$

Before we determine the series coefficients, let us consider the finiteness and continuity conditions that were noted at the beginning of the development. Finiteness was addressed by using only the spherical Bessel function for the radial dependence. This condition also led to truncation of the Legendre polynomial, without which the series would have diverged at $\psi=0$ and $\psi=\pi$. The fact that the preceding series represents a continuous pressure field results identically from the axisymmetry property. However, there is another finiteness condition that applies to the particle velocity. The velocity components are found from Euler's equation to be

$$
\begin{gather*}
V_{r}=\frac{i}{\rho_{0} c} \sum_{m=0}^{\infty} C_{m} j_{m}^{\prime}(k r) P_{m}(\cos \psi)  \tag{10.2.15}\\
V_{\psi}=-\frac{i}{\rho_{0} \omega r} \sin \psi \sum_{m=0}^{\infty} C_{m} j_{m}(k r)(\sin \psi) P_{m}^{\prime}(\cos \psi)
\end{gather*}
$$

There is nothing irregular about these expressions, but that is only because $P_{m}^{\prime}$ $(\cos \psi) \equiv 0$ at $\psi=0$ and $\psi=\pi$. Suppose this was not true. Points on opposite sides of the polar axis correspond to $\psi=\varepsilon$ at azimuthal positions $\theta$ and at $\theta+\pi$. In the limit as $\varepsilon \rightarrow 0$, the value of $V_{\psi}$ would be the same for both points, but the direction of $\bar{e}_{\psi}$ is opposite. This would correspond to an infinite velocity gradient. According to the continuity equation, such a condition corresponds to an infinite pressure.

Now that we know that the spherical harmonic series satisfies all finiteness and continuity conditions, we may proceed to evaluate the coefficients. They are determined by matching the particle velocity to the vibration of the spherical wall. The normal to the wall is $-\bar{e}_{r}$. Our convention is that the surface normal points into the fluid, so we consider positive values of the wall's normal velocity to be inward. Thus, we require that

$$
\begin{equation*}
v_{r}=-\operatorname{Re}\left(\hat{V}_{S}(\psi) e^{i \omega t}\right) \text { at } r=a \tag{10.2.16}
\end{equation*}
$$

The amplitude $V_{S}$ may be complex, in which case it represents a motion of the surface in which not all points vibrate in-phase.

Our task now is to determine the coefficients $C_{m}$ of the pressure series by satisfying the velocity boundary condition. This is done by using a Legendre function series to describe $V_{s}(\psi)$. Doing so yields a representation that is consistent with the above expression for $V_{r}$ of the fluid. Thus, we write

$$
\begin{equation*}
\hat{V}_{S}=\sum_{m=0}^{\infty} V_{m} P_{m}(\cos \psi) \tag{10.2.17}
\end{equation*}
$$

An expression for the coefficients of this series results from application of the orthogonality property of Legendre functions, Eq.(7.1.16). To that end, we multiply the series by $P_{j}(\cos \psi) \sin \psi$, where $j$ is a specific index, and then integrate over the full range of $\psi$. Changing the index $j$ back to $m$ gives

$$
\begin{equation*}
V_{m}=\frac{2 m+1}{2} \int_{0}^{\pi} \hat{V}_{S} P_{m}(\cos \psi)(\sin \psi) d \psi \tag{10.2.18}
\end{equation*}
$$

We equate the radial velocity at $r=a$, which is obtained from the first of Eq. (10.2.15) to the imposed velocity in Eq. (10.2.17). We could apply the orthogonality property, but a briefer approach argues that the Legendre functions constitute a complete set of linearly independent basis functions. Therefore, it must be that the series coefficients match, which leads to

$$
\begin{equation*}
\frac{i}{\rho_{0} c} C_{m} j_{m}^{\prime}(k a)=V_{m} \tag{10.2.19}
\end{equation*}
$$

The resulting description of the pressure is

$$
\begin{equation*}
P=\rho_{0} c \sum_{m=0}^{\infty} V_{m} \frac{j_{m}(k r)}{i j_{m}^{\prime}(k a)} P_{m}(\cos \psi) \tag{10.2.20}
\end{equation*}
$$

Earlier we identified resonances as singularities of a frequency-domain description. The singularities of the preceding expression correspond to zeros of the denominator, so we have identified that the $m$ th resonance occurs when

$$
\begin{equation*}
\left.j_{m}^{\prime}(x)\right|_{x=k a} \equiv \frac{m}{k a} j_{m}(k a)-j_{m+1}(k a)=0 \tag{10.2.21}
\end{equation*}
$$

This condition can occur for any spherical harmonic. Furthermore, the spherical functions are oscillatory, so there are multiple roots for a specific $m$. Thus, there is a double infinity of $k a$ values at which the cavity will resonate.

Proper evaluation of the series requires that an adequate number of terms be included. Resonances are an important factor. For a selected value of $k a$, there are values of $m$ for which the resonance condition is close to being satisfied. However, $j_{m}(k a(r / a)) / j_{m}^{\prime}(k a)$ in Eq. (10.2.20) decreases monotonically with increasing $m$, provided that $m$ is substantially larger than $k a$ and $r / a \leq 1$. Thus, series truncation at a value of $m$ that is a multiple of $k a$, such as $\max (m)=4 k a$, should be adequate. Another aspect of series truncation is the behavior of the $V_{m}$ coefficients. At some value of $m$, these values will decrease because the corresponding Legendre function series represents a nonsingular $V_{s}$ function. Usually, truncation based on the behavior of the Bessel functions will assure inclusion of a sufficient number of terms.

In addition to the acoustic pressure, power transfer is an important physical aspect. The instantaneous radial intensity at a point on the wall is

$$
\begin{align*}
I_{r} & =\left.\left(P V_{s} e^{2 i \omega t}+P V_{s}^{*}\right)\right|_{r=a}+\text { c.c. } \\
& =\rho_{0} c \sum_{j=0}^{\infty} \sum_{m=0}^{\infty}\left(\frac{V_{j} V_{m} e^{2 i \omega t}+V_{j} V_{m}^{*}}{i}+\text { c.c. }\right) \frac{j_{m}(k a)}{j_{m}^{\prime}(k a)} P_{j}(\cos \psi) P_{m}(\cos \psi) \tag{10.2.22}
\end{align*}
$$

According to this expression, at any point on the wall power might be flowing out of, or into, the wall. The power transfer across a patch surrounding this point might have a nonzero mean value. The total power flow is obtained by integrating this expression over the entire wall. An element of area for this axisymmetric situation is $d \mathcal{S}=2 \pi a^{2}(\sin \psi) d \psi$. Orthogonality of the Legendre functions filters out the cross terms. In addition, $V_{m} V_{m}^{*}$ is a real quantity, so the complex conjugate part cancels this term. The result is that the instantaneous power flow is

$$
\begin{equation*}
\mathcal{P}=4 \pi a^{2} \rho_{0} c \sum_{m=0}^{\infty}\left(\frac{V_{m}^{2}}{i(m+1)} e^{2 i \omega t}+\text { c.c. }\right) \frac{j_{m}(k a)}{j_{m}^{\prime}(k a)} \tag{10.2.23}
\end{equation*}
$$

From this, we conclude that the average total power transfer is zero. Each term fluctuates about zero at twice the excitation frequency, but all terms are not necessarily in-phase, so there might be no instant at which $\mathcal{P}=0$.

EXAMPLE 10.3 Consider a plane wave that is external to a spherical container. Suppose it induces a radial velocity of the wall that is proportional to the external pressure. Evaluate the pressure field within the container at $k a=2$ and $k a=5$.

## Significance

The notion that the wall velocity is proportional to the external pressure is a vast simplification. For one, it ignores the fact that the presence of the sphere induces an exterior scattered pressure field. Nevertheless, the stated conditions serve to create a situation from which we may learn much. The analysis will also help us sharpen our computational skills.

## Solution

The direction in which the plane wave propagates defines the symmetry of the field, so we let this direction be parallel to the $z$ axis. When the origin of $x y z$ is placed at the center of the cavity, the complex pressure amplitude of the incident wave is $P_{0} \exp \left(-i k \bar{x} \cdot \bar{e}_{z}\right)=P_{0} \exp (-i k z)$. For a point on the surface, $z=a \cos \psi$, so we begin by setting $V_{s}=V_{0} \exp (-i k a \cos \psi)$.

The velocity coefficients for the Legendre series representation of $V_{s}$ are

$$
V_{m}=\frac{2 m+1}{2} V_{0} \int_{0}^{\pi} e^{-i k a \cos \psi} P_{m}(\cos \psi) \sin (\psi) d \psi
$$

In general, we cannot expect that a formula for the integral is available, but it is for this case. ${ }^{2}$ The result is

$$
V_{m}=(2 m+1) e^{-m \pi i / 2} j_{m}(k a)
$$

If this formula were not available, we would evaluate the coefficients numerically. The MATLAB procedure would entail setting up a loop ranging from $m=0$ to $m=$ $\max (m)$. For each $m$, the integrand would be evaluated with an anonymous function that depends on $\eta$, such as Coeff_int=@ (eta) exp (-1i*ka*eta).*leg (m,eta), where leg.m was described in Example 7.1. The quadl function would yield the coefficient according to $V(m+1)=0.5^{*}(2 * m+1) * q u a d l$ (Coeff_int, -1,1,1e-6);

The $V_{m}$ values for $k a=2$ and $k a=5$ are plotted in Fig. 1. The largest values are centered on a harmonic number that is slightly larger than $k a$. Beyond that maximum, the coefficients fall off rapidly. According to the above formula, $V_{m}$ is a real quantity if $m$ is even, and it is imaginary if $m$ is odd. To understand why this is so, consider the integrand in conjunction with writing the complex exponential as $\cos (k a \cos \psi)-i \sin (k a \cos \psi)$. With respect to $\psi=\pi / 2$, the first term is an even function, whereas the second is an odd function. At the same time, $P_{m}(\cos \psi) \sin \psi$ is even with respect to $\pi / 2$ for even $m$ and odd for odd $m$. The integral from 0 to $\pi$ of the product of even or odd functions is nonzero, whereas the integral of the product of an even and an odd function is zero. This parity behavior will be seen to have an interesting consequence in regard to the pressure distribution.


Figure 1.

Figure 1 demonstrates that the $V_{m}$ values are negligible well before $M=\max$ $(m)=4 k a$, which is the suggested rule to truncate the series. The manner in which the pressure field is to be displayed affects the algorithm that is used to evaluate the series. We will construct contours of equal $\operatorname{Re}(P(\bar{x}))$ and $\operatorname{Im}(P(\bar{x}))$. Doing so entails evaluating the pressure at a set of grid points in axisymmetric spherical coordinates and then calling on the graphics capabilities of our chosen mathematical software. Typically, a contouring graphics routine expects the position coordinates to be Cartesian. Thus, we define a set of $r$ and $\psi$ values at equal increments, $r_{n} / a=$ $n \Delta_{r}$ and $\psi_{\ell}=\ell \Delta_{\psi}$. The corresponding Cartesian coordinates of the grid points are

[^34]$z_{n, \ell} / a=n \Delta_{r} \cos \left(\ell \Delta_{\psi}\right)$ and $x_{n, \ell} / a=n \Delta_{r} \sin \left(\ell \Delta_{\psi}\right)$. (Because of axisymmetry, it is adequate to view the field in the $x z$ plane.)

Equation (10.2.20) describes $P$ at each grid point. The matrix algorithm for evaluating a modal series at any points is suitable for the present situation. Thus, we define a diagonal matrix [ $K$ ] whose elements are $K_{m, m}=V_{m} / j_{m}^{\prime}(k a)$, a rectangular matrix $[E]$ whose elements are $E_{n, m}=j_{m}\left(k r_{n}\right)$, and a rectangular matrix [ $F$ ] whose elements are $F_{m, \ell}=P_{m}\left(\cos \psi_{\ell}\right)$. The pressure at all grid points is found by computing

$$
[P]=\frac{\rho_{0} c}{i}[E][K][F]
$$

Figure 2 shows the contours of $\operatorname{Re}(P)$ and $\operatorname{Im}(P)$ for $k a=2$. The distribution of $\operatorname{Re}(P)$ is antisymmetric with respect to the plane $\psi=\pi / 2$, and the distribution of $\operatorname{Im}(P)$ is symmetric with respect to that plane. This behavior is a consequence of the real/imaginary property of the $V_{m}$ coefficients. Both $[E]$ and $[F]$ are real, and [ $K$ ] is complex solely because its elements are proportional to the $V_{m}$ coefficients. For even $m, V_{m}$ is real and $P_{m}(\cos \psi)$ is even with respect to $\psi=\pi / 2$. Division by $i$ results in imaginary quantities that are even. The converse occurs for odd $m$, because $V_{m}$ is purely imaginary and $P_{m}(\cos \psi)$ is odd with respect to $\psi=\pi / 2$. Thus, the real pressure contours stem from the odd $m$ terms, whereas the imaginary contours are the contributions of the even $m$ terms.


Figure 2
Figure 3 shows the contours for $k a=5$. The even/odd properties displayed here are the same as those for Fig. 2. There is more fluctuation in the values from point to point, which is to be expected because the frequency is higher. What is unexpected is that the overall pressure level is lower. Also notable is the fact that the real and imaginary parts of $P$ are comparable, whereas $\operatorname{Re}(P)$ is much greater than $\operatorname{Im}(P)$ at $k a=2$.


Figure 3.

The magnitude differences are a manifestation of resonance. To explore this aspect, let us return to the pressure series, Eq.(10.2.19). To compensate for differences of the overall Bessel function magnitudes for different order $m$, let us rewrite that equation in terms of a relative coefficient $\left(B_{m}\right)_{\text {rel }}$, such that

$$
\begin{gathered}
\left(B_{m}\right)_{\mathrm{rel}}=V_{m} \frac{j_{m}(k a)}{i j_{m}^{\prime}(k a)} \\
P=\rho_{0} c \sum_{m=0}^{\infty}\left(B_{m}\right)_{\mathrm{rel}} \frac{j_{m}(k a(r / a))}{j_{m}(k a)} P_{m}(\cos \psi)
\end{gathered}
$$

The ratio of Bessel functions in the summation is always $\mathrm{O}(1)$, so $\left(B_{m}\right)_{\text {rel }}$ is a measure of the amount each term contributes to the summation. Figure 4 indicates that the dominant contribution at $k a=2$ comes from the $m=1$ harmonic. The large value of $\left(B_{1}\right)_{\text {rel }}$ stems from $j_{1}^{\prime}(k a)$ being very small; $k a=2.0816$ is a zero of $j_{1}^{\prime}(k a)$. Because this is a term for which $m$ is odd, the complex pressure amplitude is essentially a real value everywhere. In other words, it is essentially in-phase. An in-phase pressure distribution, as well as large pressure values, is hallmarks of a cavity resonance. At $k a=5$, the primary contributors are the $m=3$ and $m=4$ spherical harmonics. The values of $\left(B_{3}\right)_{\text {rel }}$ and $\left(B_{4}\right)_{\text {rel }}$ are comparable in magnitude, which is why Fig. 3 indicates that the real and imaginary parts of $P(\bar{x})$ have similar magnitudes overall. The ratio of these parts is not constant over the field, which means that the pressure field does not consist of an in-phase oscillation. We conclude that $k a=5$ is not very close to a resonance.


To close this example, it is useful to take an overview. The excitation of the spherical cavity is a wavelike particle velocity at the wall. Nevertheless, there is no evidence of a comparable wave process in the pressure field. The field at any $k a$ will consist of a complex pressure amplitude that combines real and imaginary parts. The former is antisymmetric with respect to the plane $\psi=\pi / 2$, whereas the latter is symmetric with respect to that plane. If $k a$ is close to a resonance, one part will be dominant and the pressure will show an overall maximum magnitude. Off-resonance, the mix of real and imaginary parts results in an instantaneous pressure distribution that fluctuates throughout a period because of the different parities of the real and imaginary parts of the complex amplitude.

## Concentric Spheres

Systems in which the acoustic field is the region between concentric spheres are not common. Nevertheless, their analysis serves a number of purposes. The radius of the inner sphere is $b$, so the acoustic domain is $b<r<a$. The excitation might be applied to either the inner or outer wall as a specified harmonic radial velocity or pressure oscillation. The undriven surface might be rigid or pressure-release. In order to avoid unnecessary complications, we will not consider the alternate possibility that the undriven wall is locally reacting.

It is necessary to satisfy a boundary condition at both walls, but the center, $r=0$, is not contained in the region occupied by the fluid. This means that we cannot exclude a solution of the spherical Bessel equation that is singular at the center. We could use both types of spherical Hankel functions, $h_{m}^{(1)}(k r)$ and $h_{m}^{(2)}(k r)$, to form a spherical harmonic series. These functions, respectively, represent inward and outward propagation. However, as we saw in the previous example, the field in a cavity tends to have a standing wave attribute. Adding and subtracting the two types of Hankel functions lead to the spherical Bessel and Neumann functions, $j_{m}(k r)$ and $n_{m}(k r)$, see Eq.(7.1.31). A spherical harmonic series for the complex pressure amplitude results from multiplying each function at a specific $m$ by a different coefficient and then multiplying the sum of these terms by a Legendre function. In other words,

$$
\begin{equation*}
P=\sum_{m=0}^{\infty}\left[C_{m} j_{m}(k r)+D_{m} n_{m}(k r)\right] P_{m}(\cos \psi) \tag{10.2.25}
\end{equation*}
$$

The corresponding radial particle velocity is

$$
\begin{equation*}
V_{r}=\frac{i}{\rho_{0} c} \sum_{m=0}^{\infty}\left[C_{m} j_{m}^{\prime}(k r)+D_{m} n_{m}^{\prime}(k r)\right] P_{m}(\cos \psi) \tag{10.2.26}
\end{equation*}
$$

The restriction to the passive boundary being rigid or pressure-release, accompanied by allowance that the driven sphere is either the inner or outer boundary, admits four possible sets of boundary conditions. We will address the case where the inner sphere is rigid and the outer sphere is driven. The modifications required to address the other cases will be evident. We let $V_{a}(\psi)$ be the complex amplitude of the radial velocity at $r=a$. Our convention takes a wall velocity to be positive if it points into the fluid, so $V_{a}>0$ represents a velocity that is in the direction of $-\bar{e}_{r}$. Thus, velocity continuity at the boundaries requires that

$$
\begin{gather*}
\sum_{m=0}^{\infty}\left[C_{m} j_{m}^{\prime}(k a)+D_{m} n_{m}^{\prime}(k a)\right] P_{m}(\cos \psi)=i \rho_{0} c V_{a}(\psi)  \tag{10.2.27}\\
\sum_{m=0}^{\infty}\left[C_{m} j_{m}^{\prime}(k b)+D_{m} n_{m}^{\prime}(k b)\right] P_{m}(\cos \psi)=0
\end{gather*}
$$

As we did for the case of a full spherical cavity, we expand $V_{a}$ in a Legendre function series,

$$
\begin{equation*}
V_{a}=\sum_{m=0}^{\infty} V_{m} P_{m}(\cos \psi) \tag{10.2.28}
\end{equation*}
$$

The series coefficients are

$$
\begin{equation*}
V_{m}=\frac{2 m+1}{2} \int_{0}^{\pi} V_{a} P_{m}(\cos \psi)(\sin \psi) d \psi \tag{10.2.29}
\end{equation*}
$$

The Legendre function series for $V_{a}$ is substituted into the velocity boundary conditions, Eq. (10.2.27). Application of the orthogonality properties of the Legendre functions leads to

$$
\begin{gather*}
C_{m} j_{m}^{\prime}(k a)+D_{m} n_{m}^{\prime}(k a)=i \rho_{0} c V_{m}  \tag{10.2.30}\\
C_{m} j_{m}^{\prime}(k b)+D_{m} n_{m}^{\prime}(k b)=0
\end{gather*}
$$

Solution of these equations yields

$$
\begin{align*}
C_{m} & =i \rho_{0} c \frac{n_{m}^{\prime}(k b)}{\Delta(k a, k b)} V_{m} \\
D_{m} & =-i \rho_{0} c \frac{j_{m}^{\prime}(k b)}{\Delta(k a, k b)} V_{m} \tag{10.2.31}
\end{align*}
$$

where

$$
\begin{equation*}
\Delta_{m}(k a, k b)=j_{m}^{\prime}(k a) n_{m}^{\prime}(k b)-j_{m}^{\prime}(k b) n_{m}^{\prime}(k a) \tag{10.2.32}
\end{equation*}
$$

Resonances correspond to singular values of a $C_{m}$ and $D_{m}$. They occur when $\Delta_{m}(k a, k b)=0$. This condition occurs for each harmonic. As was true for the complete spherical cavity, the oscillatory nature of the Bessel and Neumann functions causes $\Delta_{m}$ to have zeroes at an infinite number of $k$ values. Because there are an infinite number of harmonics, there are a doubly infinite set of resonances.

EXAMPLE 10.4 A concentric inner sphere in a spherical cavity executes a radially symmetric vibration, so that $v_{r}=v_{0} \cos (\omega t)$ over its entire surface. The outer sphere is rigid. (a) Evaluate the radial dependence of the pressure for $k a=2,8$, and 32 in the cases where $b=0.5$ and $b=0.1$. (b) Evaluate the dependence of $|P|$ on $k a$ at $r=a, r=(a+b) / 2$, and $r=b$ for the case where $b=0.1 a$. (c) Consider the limit as $b / a \rightarrow 0$ with the source's volume velocity, $\hat{Q}=4 \pi b^{2} v_{0}$, held constant. Derive an expression for the pressure within the sphere.

## Significance

The analysis will connect the present development with our investigations of radially symmetric waves and sources. The outcome will be enhanced understanding of acoustic behavior of cavities.

## Solution

Radially symmetric situations correspond to the $m=0$ spherical harmonic. Although the formulas derived in this section are quite general, the simplicity of the $m=0$ relations allow us to formulate the solution in terms of elementary functions. For $m=0$, we have $j_{0}(x)=\sin (x) / x$ and $n_{0}(x)=-\cos (x) / x$. Thus, the $m=0$ spherical harmonic for the region between concentric spheres is

$$
\begin{equation*}
P=\frac{1}{k r}\left[C_{0} \sin (k r)-D_{0} \cos (k r)\right] \tag{1}
\end{equation*}
$$

The complex radial velocity on the inner sphere is $v_{0}$, and the outer sphere is stationary, so the boundary conditions are

$$
\left.\frac{d P}{d r}\right|_{r=b}=-i \rho_{0} \omega v_{0},\left.\quad \frac{d P}{d r}\right|_{r=a}=0
$$

Substitution of Eq. (1) into these conditions gives

$$
\begin{gathered}
\frac{1}{k b^{2}}[k b \cos (k b)-\sin (k b)] C_{0}+[k b \sin (k b)+\cos (k b)] D_{0}=-i \rho_{0} \omega v_{0} \\
\frac{1}{k a^{2}}[k a \cos (k a)-\sin (k a)] C_{0}+[k a \sin (k a)+\cos (k a)] D_{0}=0
\end{gathered}
$$

Introduction of the identities for the sine and cosine of the difference of two angles reduces the solution of these equations to

$$
\begin{align*}
C_{0} & =-i \omega \rho_{0} k b^{2} v_{0} \frac{k a \sin (k a)+\cos (k a)}{\Delta_{0}(k a, b / a)} \\
D_{0} & =i \omega \rho_{0} k b^{2} v_{0} \frac{k a \cos (k a)-\sin (k a)}{\Delta_{0}(k a, b / a)} \tag{2}
\end{align*}
$$

where

$$
\begin{equation*}
\Delta_{0}(k a, b / a)=\left[\frac{b}{a}(k a)^{2}+1\right] \sin \left(k a\left(1-\frac{b}{a}\right)\right)-k a\left(1-\frac{b}{a}\right) \cos \left(k a\left(1-\frac{b}{a}\right)\right) \tag{3}
\end{equation*}
$$

An expression for $P$ as a function of $r$ results from substitution into Eq. (1) of the coefficients in Eq. (2). Introduction of $\omega \equiv k c$ and $b \equiv a(b / a)$ yields a nondimensional form that exhibits the fundamental parametric dependencies,

$$
\begin{equation*}
P=-i \rho_{0} c v_{0} \frac{(b / a)^{2}}{r / a} \frac{k a}{\Delta_{0}(k a, b / a)}[k a \cos (k a(r / a-1))+\sin (k a(r / a-1))] \tag{4}
\end{equation*}
$$

According to this expression, the complex pressure amplitude everywhere is purely imaginary. This means that the pressure is proportional to $\sin (\omega t)$. The radial velocity of the inner sphere is a cosine, so the pressure is in-phase with the surface acceleration.

To obtain the radial pressure distributions called for in Part (a), we evaluate Eq. (4) for fixed $b / a$ and $k a$ over $b / a<r / a<1$. The radial profiles of pressure in Fig. 1 are typical. The overall magnitude of each profile is set by the factors preceding the bracketed term in Eq. (4). The value of $\Delta_{0}$ is especially important in this regard. In most cases, the increase of pressure with decreasing radial distance is evident. However, there is little dependence on $r / a$ at low $k a$, regardless of the value of $b / a$. This behavior is consistent with our general expectation for low-frequency oscillations in a cavity. At fixed $b / a$, the radial interval over which the pressure fluctuates decreases with increasing frequency. Note that $d P / d r$ is zero at $r=a$, as required by that boundary being rigid.


Figure 1.

Frequency response functions offer a different view. Figure 2 shows the result at three locations for the case $b=a / 10$. The peaks are resonances at which $\Delta_{0}=0$. Resonance is a system feature, so they are manifested at each of the radial distances. The nulls are frequencies at which the numerator in Eq. (4) vanishes for each $r$. For a given value of $b / a$, the occurrence of this condition depends on both $k a$ and $r / a$, so the nulls are not aligned. The frequencies at which nulls occur at $r=(a+b) / 2$ become increasingly close to the resonances as the frequency increases. This behavior is a consequence of $b / a$ being quite small in the present example. When $k a \gg 1, b / a \ll 1$, and $r=(a+b) / 2$, the numerator in Eq. (4) for $r=(a+b) / 2$ is approximately proportional to $k a \cos (k a / 2)$ and $\Delta_{0}$ is approximately proportional to $(k a)^{2} \sin (k a) \equiv 2(k a)^{2} \sin (k a / 2) \cos (k a / 2)$. Resonances, at which $\Delta_{0}=0$, occur approximately at $\sin (k a / 2)=0$ and $\cos (k a / 2)=0$. The latter are approximately the same as the zeros of the numerator. When all terms are included, the values of $k a$ at which the numerator for $r=(a+b) / 2$ is zero are close to some frequencies at which $\Delta_{0}=0$.


Figure 2.

To evaluate the limit of a very small sphere, we use $v_{0}=\hat{Q} /\left(4 \pi b^{2}\right)$ in order to remove $v_{0}$ from Eq. (4). Doing so gives

$$
\begin{equation*}
P=-i \rho_{0} c \frac{\hat{Q}}{4 \pi a^{2}} \frac{a}{r} \frac{k a}{\Delta_{0}(k a, b / a)}[k a \cos (k a(r / a-1))+\sin (k a(r / a-1))] \tag{5}
\end{equation*}
$$

This expression does not contain $b$, so it is valid in the limit as $b \rightarrow 0$. Note that the limiting form requires that the inner sphere be present, so $r=0$ is excluded. In effect, Eq. (5) is the solution for a spherical cavity with a point source at its center. The pressure is singular at the source.

### 10.2.3 Cylindrical Enclosures

The mathematical representation of the pressure field in a rectangular cavity does not depend on which wall vibrates. For a spherical cavity, there is only one mathematical form. In contrast, the mathematical nature of the field in a closed cylinder depends on whether the excitation is an imposed vibration of an end cap or the cylindrical surface. We will explore both possibilities. For the sake of simplicity and brevity, only axisymmetric situations are addressed. More general situations are addressed by a simple adaptation of the present results. This is so because a Fourier series description of an arbitrary $\theta$ dependence uncouples due to orthogonality of the circumferential harmonics.

## Vibrating End Wall

We consider the end $x=0$ to execute a specified vibration in an axisymmetric pattern, so that $V_{x}=V_{s}(R)$. The end $x=L$ is a locally reacting surface with specific impedance $\zeta_{L}$. The transverse mode functions within a cylinder were derived in Sect.9.3.3. The axisymmetric modes are

$$
\begin{equation*}
\Phi_{0, j}(R)=B_{0, j} J_{0}\left(\eta_{0, j} \frac{R}{a}\right), \quad \mu_{0, j}=\frac{\eta_{0, j}}{a} \tag{10.2.33}
\end{equation*}
$$

The $B_{0, j}$ values in Eq. (10.2.33) are selected to normalize the mode functions according to Eq. (9.3.67).

The $\eta_{0, j}$ quantities are eigenvalues. The characteristic equation from which they are obtained results from satisfying velocity continuity at the cylindrical wall, where the specific impedance is $\zeta_{a}$. Equation (9.3.57) is the result that was derived in Chap. 9. Because only the $n=0$ harmonic is being considered, that expression may be simplified by exploiting the identity that $J_{0}^{\prime}\left(\eta_{0, j}\right) \equiv-J_{1}\left(\eta_{0, j}\right)$. The characteristic equation correspondingly may be written as

$$
\begin{equation*}
J_{0}\left(\eta_{0, j}\right)+\frac{i \zeta_{a}}{k a} \eta_{0, j} J_{1}\left(\eta_{0, j}\right)=0 \tag{10.2.34}
\end{equation*}
$$

Evaluation of the modes was discussed in Sect.9.3.3. The eigenvalues are sequenced in ascending order of their real part. In the case where the wall is rigid, $j=0$ is the smallest eigenvalue, $\eta_{0,0}=0$. This is the plane wave mode. It cannot exist if the cylindrical wall is not rigid, because a pressure in the wall would induce a particle velocity in the $\bar{e}_{R}$ direction, whereas the velocity in a plane wave is solely in the $x$ direction.

If the cylindrical wall impedance is a compliance, that is, if $\zeta_{a}$ is a negative imaginary value, then there is one imaginary eigenvalue. That mode is designated as $j=0$. The modal series that we will develop will assume that a mode for $j=0$ exists. If it does not, then the associated modal series coefficient should be set to zero.

In the axial direction, we are presented with the same situation as that for a rectangular enclosure, in that each transverse mode is associated with a pair of waves that propagate forward and back in the $x$ direction. Hence, a modal series will have the same form as Eq. (10.2.2), except that a single summation is used because only the $n=0$ azimuthal harmonic is excited. The representation of the pressure field is

$$
\begin{equation*}
P=\sum_{j=0}^{\infty}\left[\alpha_{0, j} e^{-i \kappa_{0, j} x}+\beta_{0, j} e^{+i \kappa_{0, j} x}\right] \Phi_{0, j}(R) \tag{10.2.35}
\end{equation*}
$$

The axial wavenumbers are the values that are set by the requirement that each term independently satisfies the Helmholtz equation, which gives

$$
\begin{equation*}
\kappa_{0, j}=\left[k^{2}-\left(\eta_{0, j} / a\right)\right]^{1 / 2} \tag{10.2.36}
\end{equation*}
$$

The series accounts for waves that propagate in both directions, so we may retain the requirement that the square root be selected such that $\operatorname{Re}\left(\kappa_{0, j}\right) \geq 0$ and $\operatorname{Im}$ $\left(\kappa_{0, j}\right) \leq 0$. In turn, this requires that the $\eta_{0, j}$ values be in the first or second quadrant of the complex plane, which includes values on the positive real axis.

The steps leading to the determination of the modal coefficients parallel those for analysis of a rectangular cavity. The boundary conditions at the driven end and the opposite passive end are like those in Eq.(10.2.4). Consequently, the results
may be derived from those for the rectangular case by modifying any variables associated with the transverse behavior to match the circular geometry. The analog of Eq.(10.2.7) is an inner product over the circular cross section. A differential element of area for an axisymmetric integrand may be taken to be $d \mathcal{A}=2 \pi R d R$, and the cross-sectional area is $\pi a^{2}$, so the modal velocity coefficients are

$$
\begin{equation*}
V_{0, j}=\frac{1}{\pi a^{2}} \int_{0}^{a} \Phi_{0, j}(R) V_{s}(R) 2 \pi R d R \tag{10.2.37}
\end{equation*}
$$

Aside from the fact that the transverse modes are those for a cylindrical waveguide, the pressure field is the same as Eq. (10.2.10),

$$
\begin{align*}
P= & \rho_{0} c \sum_{j=1}^{\infty} \frac{k V_{0, j}}{\kappa_{0, j} \Delta\left(\kappa_{0, j}\right)}\left[\left(1+\zeta_{L} \frac{\kappa_{0, j}}{k}\right) e^{i \kappa_{0, j}(L-x)}\right.  \tag{10.2.38}\\
& \left.-\left(1-\zeta_{L} \frac{\kappa_{0, j}}{k}\right) e^{-i \kappa_{0, j}(L-x)}\right] \Phi_{0, j}(R)
\end{align*}
$$

The quantity $\Delta\left(\kappa_{0, j}\right)$ in the denominator is the same as the determinant in Eq.(10.2.9), except that the axial wavenumbers are those for the cylinder,

$$
\begin{equation*}
\Delta\left(\kappa_{0, j}\right)=\left(1+\zeta_{L} \frac{\kappa_{0, j}}{k}\right) e^{i \kappa_{0, j} L}+\left(1-\zeta_{L} \frac{\kappa_{0, j}}{k}\right) e^{-i \kappa_{0, j} L} \tag{10.2.39}
\end{equation*}
$$

As was true for a rectangular cavity, resonances correspond to $\Delta=0$. This condition can occur only if $\operatorname{Re}\left(\zeta_{L}\right)=0$ and $\kappa_{0, j}$ is real. This feature is evident in the alternate form of $\Delta$ obtained by application of Euler's identity. It is

$$
\begin{equation*}
\Delta\left(\kappa_{0, j}\right)=2\left[\cos \left(\kappa_{0, j} L\right)+i \zeta_{L} \frac{\kappa_{0, n}}{k} \sin \left(\kappa_{0, j} L\right)\right] \tag{10.2.40}
\end{equation*}
$$

The procedure by which the natural frequencies may be analyzed when $\zeta_{a}$ or $\zeta_{L}$ is reactive is essentially the same as that discussed in regard to Eq.(10.2.13). The primary difference is that the restriction to axisymmetric modes means that the transverse mode function depends on a single eigenvalue $\eta_{0, j}$, and the characteristic equation for that value features Bessel functions, rather than harmonic functions. The natural frequencies are found by a frequency sweep ranging over $k$ at fixed $j$. The value $\eta_{0, j}$ is combined with the set of roots $\kappa_{0, j}^{(m)}$ for which $\Delta\left(\kappa_{0, j}\right)=0$ to evaluate tentative natural frequencies according to

$$
\begin{equation*}
\left(k_{\text {nat }}\right)_{0, j}=\left[\left(\kappa_{0, j}^{(m}\right)^{2}+\left(\frac{\eta_{0, j}}{a}\right)\right]^{1 / 2} \tag{10.2.41}
\end{equation*}
$$

Any of these values that is very close to the value of $k$ for the computations marks a natural frequency, $\omega_{0, j}^{(m)}=\left(k_{\text {nat }}\right)_{0, j} / c$.

EXAMPLE 10.5 The concept of an impedance tube developed in Sect. 3.2.4 is that measured properties of the field within a closed waveguide can be used to identify the impedance of a material placed at one end of the tube. The concept was derived on the basis of the field being planar in the transverse direction, but that is not possible in reality. This is so because a planar waveguide mode exists only if the sidewalls are rigid, which cannot be for actual materials. Consider the impedance tube in Example 3.6. As described there, the tube is a cylinder whose diameter is 120 mm , and $L=2 \mathrm{~m}$, The pressure field is generated at $x=0$ by a vibrating piston whose diameter equals that of the tube. The specific impedance of the material that covers the end at $x=L$ was $\zeta_{L}=2.763+1.168 i$. (This is the value used to generate the original data.) The fluid within the impedance tube is water, $\rho_{0}=1000 \mathrm{~kg} / \mathrm{m}^{3}, c=1480 \mathrm{~m} / \mathrm{s}$. Determine and graph the meansquared pressure distribution along the centerline when the cylindrical wall is compliant, with $\zeta_{a}=-5 i$. Compare the result to the axial distribution obtained from an idealized rigid wall model. Frequencies to consider are 600 Hz , as stated in the original problem, and 6 kHz .

## Significance

Flexibility of the sidewall is a primary issue for the design of impedance tubes. From an instructional standpoint, this example serves to connect the multidimensional phenomena to our earlier studies of plane waves.

## Solution

The first task is evaluation of the eigenvalues of the transverse mode functions. The cylindrical wall's impedance is a pure compliance. Consequently, the real eigenvalues $\eta_{0, j}$ are the roots of Eq. (10.2.34). There also is one imaginary eigenvalue, $\eta_{0, \text { im }}=$ $i \beta a$, where $\beta$ is the root of Eq. (9.3.60) with $n=0$. This is the type of mode for which $j=0$ is reserved. The first few eigenvalues for each frequency and corresponding axial wavenumbers are tabulated below. The oscillatory nature of Bessel functions is such that eigenvalues for the higher modes are well approximated by incrementing $\eta_{0,4}$ by multiples of $\pi$.

|  |  | $j$ |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $f(\mathrm{~Hz})$ |  | im | 1 | 2 | 3 | 4 |
| 600 | $\eta_{0, j}$ | $0.5022 i$ | 3.7997 | 6.9981 | 10.1614 | 13.3145 |
|  | $\kappa_{0, j}$ | 0.7911 | $-3.7502 i$ | $-6.9714 i$ | $-10.1430 i$ | $-13.3005 i$ |
| 6000 | $\eta_{0, j}$ | $1.8306 i$ | 3.5107 | 6.8409 | 10.0532 | 13.2319 |
|  | $\kappa_{0, j}$ | 6.3816 | 5.0048 | $-3.0700 i$ | $-7.9808 i$ | $-11.7349 i$ |

At 600 Hz , for which $k a=0.611$, all modes except the one associated with $\eta_{0, \text { im }}$ are cut off. Raising the frequency to 6 kHz only adds one mode to the set of real axial wavenumbers.

The $B_{0, j}$ coefficients that normalize the transverse mode functions are described in Eqs. (9.3.67) and (9.3.66). The result for $n=0$ is

$$
B_{0, j}=\frac{\eta_{0, j}}{J_{0}\left(\eta_{0, j}\right)}\left[\left(\frac{k a}{\zeta_{a}}\right)^{2}+\left(\eta_{0, j}\right)^{2}\right]^{-1 / 2}
$$

To evaluate the coefficients of the modal series for the velocity excitation, we use Eq. (10.2.37) with $V_{s}$ set to a constant value $V_{\mathrm{p}}$ for a vibrating piston. The resulting integral is a standard one, ${ }^{3}$

$$
\frac{V_{0, j}}{V_{\mathrm{p}}}=\frac{1}{\pi a^{2}} \int_{0}^{a} B_{0, j} J_{0}\left(\eta_{0, j} R\right) 2 \pi R d R=\frac{2 B_{0, j}}{\eta_{0, j}} J_{1}\left(\eta_{0, j}\right)
$$

This expression indicates that the $V_{0, j}$ values decrease with increasing $j$, both as a consequence of the overall decrease of Bessel functions as their argument increases and as a consequence of the presence of $\eta_{0, j}$ in the denominator. Figure 1 describes these values, with $V_{0, \text { im }}$ assigned to $j=0$. Beyond $j=8$, the coefficients are less than $1 \%$ of the largest at both frequencies. Thus, the modal series will omit terms for which $j>8$. (This actually is many more terms than necessary for series convergence, because all modes for $j>2$ at both frequencies are cut off. The corresponding eigenvalues are negative imaginary, which will lead to $\Delta(k, j)$ being very large.)


Figure 1.
It is desired to evaluate the pressure at $R=0$, where the zero-order Bessel function is one. Therefore, the on-axis pressure is obtained by setting $\Phi_{0, j}=B_{0, j}$ in Eq. (10.2.38). Computation of the field at a set of points $x_{m}$ along this axis may be performed by the matrix algorithm used previously. The $V_{0, j} / V_{\mathrm{p}}$ values form column vector $\{V\}$. A rectangular array $[\Phi]$ holds the terms in Eq. (10.2.38) that depend on axial position, that is,

$$
\Phi_{0, n}=\left(1+\zeta_{L} \frac{\kappa_{0, n}}{k}\right) e^{i \kappa_{0, n}\left(L-x_{m}\right)}-\left(1-\zeta_{L} \frac{\kappa_{0, n}}{k}\right) e^{-i \kappa_{0, n}\left(L-x_{m}\right)}
$$

Then, the complex pressure amplitudes at the successive locations are found by evaluating

$$
\left\{\frac{P}{\rho_{0} c V_{\mathrm{p}}}\right\}=[\Phi]\left[\frac{k}{\kappa_{0, n} \Delta\left(\kappa_{0, j}\right)}\right]_{\mathrm{diag}}\{V\}
$$

[^35]where $\Delta\left(\kappa_{0, j}\right)$ is described by Eq. (10.2.39). This computation is performed for both frequencies.

The pressure for planar waves in a closed waveguide (with rigid walls) is given in Eq. (3.2.21). It will be easier to interpret the results if this expression is written in terms of $\zeta_{L}$ rather than the reflection coefficient at $z=L$. Such a form is obtained by multiplying that expression by $\zeta_{L}+1$, which leads to

$$
\begin{aligned}
\frac{P_{\mathrm{rigid}}(x)}{\rho_{0} c V_{\mathrm{p}}} & =\frac{\left(1+\zeta_{L}\right) e^{i k(L-x)}-\left(1-\zeta_{L}\right) e^{-i k(L-x)}}{\Delta_{\text {rigid }}} \\
\Delta_{\text {rigid }} & =\left(1+\zeta_{L}\right) e^{i k L}+\left(1-\zeta_{L}\right) e^{-i k L}
\end{aligned}
$$

The quantity used for an impedance tube calculation is the mean-squared pressure. The axial dependence of $\left(p^{2}\right)_{\mathrm{av}}$ is depicted in Fig. 2. Clearly, there is a drastic difference between the behavior at the two frequencies, as well as between the rigid waveguide model and the model that account for the compliance of the wall. To understand why this is so, let us first focus on the results at 600 Hz .


Figure 2.
We found that the zero mode, corresponding to $\eta_{0, \text { im }}$, is the only one that is not cut off at 600 Hz . For all other modes, $\kappa_{0, j}$ is a negative imaginary value. The value of $\Delta(k, j)$ for these modes is very large, which means that these modes contribute little to the modal pressure series. Indeed, the graph of $\left(p^{2}\right)_{\mathrm{av}}$ would be unchanged if only the "im" mode was used. This explains why the plot of $\left(p^{2}\right)_{\mathrm{av}}$ at 600 Hz has a sinusoidal appearance.

To explain why this plot differs for that for a rigid wall, we need to compare the $n=0$ term in Eq. (10.2.38) to the expression for $P_{\text {rigid }}$. The contribution of the "im" mode to $P /\left(\rho_{0} c V_{\mathrm{p}}\right)$ is proportional to $k / \kappa_{0, \mathrm{im}}$, which is 0.77 , and $V_{0, \mathrm{im}} / V_{\mathrm{p}}$, which is 0.9998 . Thus, the different magnitude of the modal description of $P$ and $P_{\text {rigid }}$ must stem from differences of the respective $\Delta$ values. For the "im" mode, this is $\Delta\left(\kappa_{0,0}\right)=0.9838+2.1789 i$, whereas $\Delta_{\text {rigid }}=2.9133-5.1275 i$. A resonance corresponds to $\Delta$ being zero, so we conclude that the "im" mode is closer at 600 Hz to a resonance than is the rigid model.

The fact that the pressure amplitudes are different does not necessarily invalidate application of the impedance tube concept, which relies on the maximum and minimum values of $\left(p^{2}\right)_{\mathrm{av}}$. The procedure derived in Sect.3.2.4 for extracting $\zeta_{L}$ from measured data should not be applied directly because the axial wavenumber is $\kappa_{0, \text { im }}$, rather than $k$. A comparison of the "im" term in the modal series, Eq. (10.2.38), to the above expression for $P_{\text {rigid }}$ suggests that the procedure may be used if we set the wavenumber to $\kappa_{0, \text { im }}$ and replace the specific impedance with an effective value given by

$$
\zeta_{\mathrm{eff}}=\zeta_{L} \frac{\kappa_{0, \mathrm{im}}}{k}
$$

Note that in an actual setup we could identify $\kappa_{0, \text { im }}$ by measuring the distance between adjacent maxima or minima of $\left(p^{2}\right)_{\mathrm{av}}$.

The ratio of the maximum and minimum values of $\left(p^{2}\right)_{\mathrm{av}}$ is defined as $S^{2}$ in Sect.3.2.4. The present data for 600 Hz gives $S=4.2572$. From this, we find the magnitude of the modal reflection coefficient, see Eq. (10.2.11),

$$
\left|\left(R_{L}\right)_{0, \mathrm{im}}\right|=\frac{S-1}{S+1}=0.620
$$

To find $\delta$, which is the argument of $\left(R_{L}\right)_{0, \text { im }}$, we use Eq. (3.2.57) with $k$ replaced by $\kappa_{0, i m}$. The value of $\delta$ we shall use is the average of the value obtained from each maximum or minimum. The result is $\delta=0.230$, so the reflection coefficient is $\left(R_{L}\right)_{0, \text { im }}=0.620 \exp (0.230 i)$. In turn, the effective impedance is found to be

$$
\zeta_{\mathrm{eff}}=\frac{1+\left(R_{L}\right)_{0, \mathrm{im}}}{1-\left(R_{L}\right)_{0, \mathrm{im}}}=3.4727+1.5937 i
$$

The corresponding true impedance is

$$
\zeta_{L}=\frac{k}{\kappa_{0, \mathrm{im}}} \zeta_{\mathrm{eff}}=2.6836+1.2316 i
$$

The latter value is quite close to the impedance used to generate the data.
The evaluation of $\zeta_{L}$ at 600 Hz yielded a meaningful result because the concept is based on the axial pressure distribution being sinusoidal. This is not the case at 6 kHz . We previously established that the $j=1$ mode, as well as the "im" mode, is not cut off at this frequency. Furthermore, $V_{0,1}$ is $20 \%$ of $V_{0, i m}$. The respective axial wavelengths are $2 \pi / \kappa_{0, \mathrm{im}}=0.236 \mathrm{~m}$ and $2 \pi / \kappa_{0,1}=0.301 \mathrm{~m}$. The combination of the contributions of these modes is the irregular dependence of $\left(p^{2}\right)_{\mathrm{av}}$ that appears in Fig. 2. This negates identification of $\zeta_{L}$ at 6 kHz by the scheme developed in Sect.3.2.4.

Our analysis leads to the conclusion that the given waveguide design is useful at any frequency where the "im" mode is the only one whose eigenvalue is real. As the frequency is increased from 600 Hz , the $j=1$ mode "cuts on" when $k a=\eta_{0, j}$. This frequency is 3572 Hz . For comparison, if the wall were rigid, the axial wavenumber of the $j=1$ mode, which is the lowest nonplanar mode, is real above 3760 Hz .

Increasing the stiffness corresponds to making $\operatorname{Im}\left(\zeta_{a}\right)$ a larger negative number. This causes $\eta_{0, \text { im }}$ to decrease toward zero and each $\eta_{0, j}$ for $j \geq 1$ to approach the respective value for a rigid wall. From this, we may deduce some design objectives for an impedance tube. The first nonplanar mode is cut off below $\omega_{\text {cut }}=\eta_{0,1} c / a$. Thus, $a$ should be minimized. Then, the walls should be made as rigid as possible, so that the cutoff frequencies of the nonplanar modes are close to the ideal values for a rigid wall. Although a locally reacting model is a useful convenience, any wall is part of a dynamic structure that has its own set of natural frequencies. The impedance of the boundary as seen by the fluid within the cavity is greatly reduced at these frequencies. Nevertheless, these characteristics do not alter the basic conclusions, most important of which is that usage of an impedance tube is limited to frequencies below the cutoff frequency of the second mode.

## Vibrating Cylindrical Wall

Regardless of how a cylindrical cavity is excited, a series of forced modes must feature a sum of products of a Bessel function of $\mu R$ and a complex exponential $\exp ( \pm i \kappa x)$, because such terms constitute a fundamental solution of the Helmholtz equation for a cylindrical domain. The excitation in the previous section was an imposed vibration at an end, so that waves propagated in the $x$ direction. When the excitation of a cylindrical cavity is an imposed vibration of the cylindrical boundary, a cylindrical wave propagates inward. Its arrival at the centerline generates an outward wave that cancels the singularity of the inward wave. The axial direction $x$ now is transverse to the propagation direction. Correspondingly, the $x$ dependence is set by the need to fit axial waves between the ends. In this view, the significance of $R$ and $x$ is swapped relative to that for vibration at an end.

The wall vibration $V_{a}$ is a radial velocity, positive into the fluid, that may depend arbitrarily on the axial position. An axisymmetric field occurs when $V_{a}$ is independent of $\theta$. The ends at $x=0$ and $x=L$ may be locally reacting, with specific impedances $\zeta_{0}$ and $\zeta_{L}$. Thus, we seek a solution of the Helmholtz equation that is finite at $R=0$ and satisfies impedance boundary conditions at $x=0$ and $x=L$,

$$
\begin{align*}
P & =\frac{\zeta_{0}}{\iota k} \frac{\partial P}{\partial x} \text { at } x=0 \\
P & =-\frac{\zeta_{L}}{\iota k} \frac{\partial P}{\partial x} \text { at } x=L \tag{10.2.42}
\end{align*}
$$

In addition, the (inward) particle velocity at the cylindrical wall must match the input, so it must be that

$$
\begin{equation*}
\frac{\partial P}{\partial R}=+i \omega \rho_{0} V_{s}(x) \text { at } R=a \tag{10.2.43}
\end{equation*}
$$

A comparison of the boundary conditions at the ends, Eq. (10.2.42), to those at the sidewalls of a two-dimensional waveguide, Eq. (9.2.29), shows that they only differ in the labels of parameters and coordinates. Thus, the axial mode functions here will be the same as the transverse functions in Eq. (9.2.33), with $y, H$, and $\eta_{n}$ changed to $x, L$, and $\eta_{0, j}=\kappa_{0, j} L$, respectively. These functions, therefore, are

$$
\begin{equation*}
\Phi_{0, j}(x)=B_{0, j}\left[k L \sin \left(\eta_{0, j} \frac{x}{L}\right)-i \zeta_{0} \eta_{0, j} \cos \left(\eta_{0, j} \frac{x}{L}\right)\right] \tag{10.2.44}
\end{equation*}
$$

Like the notation for excitation at an end, using a zero first subscript reminds us that this is an axisymmetric function. The second subscript is the number of the root of the characteristic equation. As always, $j=0$ is reserved for a zero or an imaginary eigenvalue. Adaptation of Eq. $(9.2 .32)$ to the current notation gives the characteristic equation for the eigenvalues $\eta_{0, j}$,

$$
\begin{equation*}
-\left[(k L)^{2}+\zeta_{0} \zeta_{L}\left(\eta_{0, j}\right)^{2}\right] \sin \left(\eta_{0, j}\right)+i\left(\zeta_{0}+\zeta_{L}\right)(k L)\left(\eta_{0, j}\right) \cos \left(\eta_{0, j}\right)=0 \tag{10.2.45}
\end{equation*}
$$

The functions $\tilde{\Phi}_{0, j}$ in Eq. (10.2.44) have a magnitude that varies substantially as the eigenvalue changes. Normalization ensures that the mode functions we use do not share that attribute. The coefficient $B_{0, j}$ is normalized consistently with Eq. (9.2.51). An inner product of the mode function is defined relative to the $x$ direction, so the $B_{0, j}$ values are evaluated according to

$$
\begin{equation*}
\left(B_{0, j}\right)^{2}=\frac{L}{\int_{0}^{L}\left(\Phi_{0, j}\right)^{2} d x} \tag{10.2.46}
\end{equation*}
$$

The other part of the forced modes for this type of excitation describes how waves propagate in the $R$ direction. This dependence must be Bessel functions because the Neumann functions are singular at $R=0$. A modal series will feature a product of a transverse mode and a Bessel function that depends on $R$. Specifically, the modal series is taken to be

$$
\begin{equation*}
P=\sum_{j=0}^{\infty} P_{0, j} J_{0}\left(\mu_{0, j} R\right) \Phi_{0, j}(x) \tag{10.2.47}
\end{equation*}
$$

The transverse wavenumber is denoted $\mu_{0, j}$, just as it was for the case of end excitation. The difference is that previously this wavenumber was obtained by solving a characteristic equation, whereas here it is set by the requirement that each term in the series satisfies the Helmholtz equation. Hence, it must be that the transverse wavenumbers are

$$
\begin{equation*}
\mu_{0, j}=\left[k^{2}-\left(\frac{\eta_{0, j}}{L}\right)^{2}\right]^{1 / 2} \tag{10.2.48}
\end{equation*}
$$

It remains to match the particle velocity to the imposed velocity of the cylinder wall. Substitution of the modal series into Eq. (10.2.43) and application of the fact that $J_{0}^{\prime}(\beta) \equiv-J_{1}(\beta)$ give

$$
\begin{equation*}
-\sum_{j=1}^{\infty} \mu_{0, j} P_{0, j} J_{1}\left(\mu_{0, j} a\right) \Phi_{0, j}(x)=i \omega \rho_{0} V_{a}(x) \tag{10.2.49}
\end{equation*}
$$

The $\Phi_{0, j}$ functions are an orthogonal set, as described by Eq. (9.2.52). We exploit this property by multiplying the preceding relation by an arbitrarily selected $\Phi_{0, n}(x)$ and then integrating the result over $0<x<L$. The result is that

$$
\begin{equation*}
P_{0, j}=-\rho_{0} c \frac{i k}{\mu_{0, j} J_{1}\left(\mu_{0, j} a\right)} V_{0, j}, \quad V_{0, j}=\frac{1}{L} \int_{0}^{L} V_{a}(x) \Phi_{0, j}(x) d x \tag{10.2.50}
\end{equation*}
$$

The pressure field corresponding to these coefficients is

$$
\begin{equation*}
P=\rho_{0} c \sum_{j=0}^{\infty} \frac{i k V_{0, j}}{\mu_{0, j} J_{1}\left(\mu_{0, j} a\right)} J_{0}\left(\mu_{0, j} R\right) \Phi_{0, j}(x) \tag{10.2.51}
\end{equation*}
$$

Resonances correspond to situations where the denominator of any term in the preceding summation is zero, that is, whenever the transverse wavenumber of the $j$ th mode is a zero of the first-order Bessel function. There are many such zeros; we will use $\ell$ to denote the root number. Thus, resonances are marked by

$$
\begin{equation*}
\mu_{0, j} a=\tilde{\mu}_{\ell} \text { where } J_{1}\left(\tilde{\mu}_{\ell}\right)=0, \ell=1,2, \ldots \tag{10.2.52}
\end{equation*}
$$

The first few roots are

$$
\begin{equation*}
\tilde{\mu}_{0}=0, \tilde{\mu}_{1}=3.8317, \tilde{\mu}_{2}=7.0156, \tilde{\mu}_{3}=10.1735 \tag{10.2.53}
\end{equation*}
$$

Higher roots are essentially integer multiples of $\pi$ greater than the last listed value. To find the natural frequencies we recall Eq. (10.2.48),

$$
\begin{equation*}
\left(k_{\mathrm{nat}}\right)_{0, j, \ell}=\left[\left(\frac{\eta_{0, j}}{L}\right)+\left(\frac{\tilde{\mu}_{\ell}}{a}\right)^{2}\right]^{1 / 2} \tag{10.2.54}
\end{equation*}
$$

An $\eta_{0, j}$ value is one of the infinite sets of roots of the characteristic equation for the axial direction, Eq. (10.2.45), but $k$ appears in that equation. Thus, unless both ends are either rigid or pressure-release, the value of $\eta_{0, j}$ depends on $k$. However, the preceding defines the value of $k$ at a natural frequency as a function of $\eta_{0, j}$. Identification of the natural frequencies in the presence of this dilemma was addressed by the discussion of Eq. (10.2.13).

EXAMPLE 10.6 A tentative design of a woofer, which is a loudspeaker that radiates low-frequency sound, consists of a cavity whose cylindrical wall is forced to undergo a uniform expansion at frequency $\omega$. One
end of the cavity is rigid, and the other is closed by a loosely tensioned membrane that well approximates a pressure-release condition. The sound generated by this speaker radiates from the vibrating membrane, so the velocity of the membrane is of interest. In particular, at low frequencies, $k a<1$, the radiated power is proportional to the volume velocity of a source. Therefore, it is desired to derive an expression for the average velocity across the membrane. Evaluate this expression for the range from 20 Hz to 1 kHz for cases where the fluid within the cavity is air and water. The diameter of the cavity is 300 mm and the length is 500 mm .

## Significance

The analysis will provide an in-depth exploration of the manner in which a vibrating cylindrical wall influences the interior field. It also will introduce some rudimentary issues for transducer design.

## Solution

Formation of the modal series begins with determination of the axial mode functions. It is convenient to let $x=0$ be the pressure-release end, so that $\zeta_{0}=0$ and $\zeta_{L}$ is infinite. We divide the characteristic equation, Eq. (10.2.45), by $\zeta_{L}$ in order to handle its infinite value, which reduces the equation to

$$
(\kappa L) \cos (\kappa L)=0
$$

The root $\kappa L=0$ leads to $\Phi_{0, j}=0$, so it is discarded. Thus, $\kappa L$ must be an odd multiple of $\pi / 2$,

$$
\begin{equation*}
\eta_{0, j}=\kappa_{0, j} L=\frac{2 j-1}{2} \pi, \quad j=1,2, \ldots \tag{1}
\end{equation*}
$$

The axial mode functions obtained from Eq. (10.2.44) are

$$
\Phi_{0, j}=B_{0, j} \sin \left(\frac{(2 j-1) \pi x}{2 L}\right)
$$

To normalize the modes we set

$$
\int_{0}^{L}\left[B_{0, j} \sin \left(\frac{(2 j-1) \pi x}{2 L}\right)\right]^{2} d x \equiv\left(B_{0, j}\right)^{2}\left(\frac{L}{2}\right)=L
$$

so the normalized modes are

$$
\begin{equation*}
\Phi_{0, j}=\sqrt{2} \sin \left(\frac{(2 j-1) \pi x}{2 L}\right) \tag{2}
\end{equation*}
$$

The transverse wavenumbers are given, in Eq. (10.2.48),

$$
\begin{equation*}
\mu_{0, j}=\frac{1}{L}\left[(k L)^{2}-\left(\frac{2 j-1}{2}\right)^{2} \pi^{2}\right]^{1 / 2} \tag{3a}
\end{equation*}
$$

The negative imaginary root is taken if $k L<(j-1 / 2) \pi$.
It is stated that the radial vibration of the cylindrical wall is uniform, which means that $V_{a}$ is a constant value. Substitution of Eq. (2) into Eq. (10.2.50) shows that the corresponding velocity coefficients are

$$
\begin{equation*}
V_{j}=\frac{V_{a}}{L} \int_{0}^{L} \Phi_{0, j} d x=\frac{2 \sqrt{2}}{(2 j-1) \pi} V_{s} \tag{4}
\end{equation*}
$$

Substitution of these quantities into the modal series, Eq. (10.2.51), gives

$$
\begin{equation*}
P=-\rho_{0} c V_{a} \sum_{j=1}^{\infty} \frac{4}{(2 j-1) \pi} \frac{i k J_{0}\left(\mu_{0, j} R\right)}{\mu_{0, j} J_{1}\left(\mu_{0, j} a\right)} \sin \left(\frac{(2 j-1) \pi x}{2 L}\right) \tag{5}
\end{equation*}
$$

If $(2 j-1) \pi>2 k L$, the mode number is above the cutoff value. Then, $\mu_{0, j}$ is negative imaginary, that is, $\mu_{0, j}=-i \hat{\mu}_{0, j}$. We may either directly input this imaginary value to the computer routine for Bessel functions, or alternatively invoke the identity that $J_{n}(-i \beta)=(-i)^{n} I_{n}(\beta)$, see Sect.9.3.3.

The quantity of interest here is the axial particle velocity at $x=0$. From Euler's equation, we have

$$
\begin{equation*}
V_{x}=-\frac{1}{i \rho_{0} \omega} \frac{\partial P}{\partial x}=V_{a} \sum_{j=1}^{\infty} \frac{2 J_{0}\left(\mu_{0, j} R\right)}{\left(\mu_{0, j} L\right) J_{1}\left(\mu_{0, j} a\right)} \cos \left(\frac{(2 j-1) \pi x}{2 L}\right) \tag{6}
\end{equation*}
$$

The volume velocity is the area integral of the particle velocity at $x=0$. A reference volume velocity is the cylindrical wall's velocity $V_{a}$ multiplied by the cross-sectional area. Thus,

$$
\frac{\tilde{Q}}{\pi a^{2} V_{a}}=\frac{1}{\pi a^{2}} \int_{0}^{a} \frac{\left.V_{x}\right|_{x=0}}{V_{2}}(2 \pi R) d R
$$

An equivalent interpretation of this expression is that it is the ratio of the average axial velocity at $x=0$ to the wall's velocity. Substitution of Eq. (6) reduces this equation to an integral of $R J_{0}\left(\mu_{0, j} R\right)$, for which a standard formula ${ }^{4}$ is available. The result is that

$$
\begin{equation*}
\frac{V_{\mathrm{av}}}{V_{a}} \equiv \frac{\tilde{Q}}{\pi a^{2} V_{a}}=\left.\frac{1}{\pi a^{2}} \int_{0}^{a} V_{x}\right|_{x=0}(2 \pi R) d R=\frac{4}{\pi} V_{a} \sum_{j=1}^{\infty} \frac{a / L}{\left(\mu_{0, j} a\right)^{2}} \tag{7}
\end{equation*}
$$

Interpretation of the frequency response graphs will be easier if we first compute the natural frequencies. The procedure for this analysis discussed in connection

[^36]with Eq.(10.2.54) is not needed because the ends are rigid and pressure-release. Consequently, the axial wavenumbers are $\eta_{0, j} / L$ are given by Eq.(1), regardless of the frequency. The resonance condition is $J_{1}\left(\mu_{0, j} a\right)=0$, whose lower roots are given in Eq. (10.2.53). Substitution of these values into Eq. (10.2.54) yields the natural frequencies. The following tabulation lists the values that occur below 1 kHz .

| Air |  |  |  | Water |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Axial \#j $\ell$ | $\tilde{\mu}_{\ell}$ | $\left(\omega_{\text {nat }}\right)_{j, \ell}(\mathrm{~Hz})$ | Axial \#j $\ell$ | $\tilde{\mu}_{0, j}\left(\omega_{\text {nat }}\right)_{j, \ell}(\mathrm{~Hz})$ |  |  |  |
| 1 | 0 | 0 | 170 | 1 | 0 | 0 | 740 |
| 2 | 0 | 0 | 510 |  |  |  |  |
| 1 | 1 | 3.8317 | 711.7 |  |  |  |  |
| 3 | 0 | 0 | 850 |  |  |  |  |
| 2 | 1 | 3.8317 | 858.9 |  |  |  |  |

Figure 1 shows the result of evaluating Eq. (7) for the average velocity. When the cylinder is filled with water, the single resonance at 740 Hz matches the tabulated natural frequency. In contrast, for the case where air is the fluid, only the natural frequencies at 170,510 , and 850 Hz are manifested as resonances. (The data was probed to ascertain that the latter is not the adjacent natural frequency of 858.9 Hz .)


Figure 1.
Do the missing resonances represent an anomaly? Let us look at the pressure on-axis at the midpoint, which is described by Fig. 2. Every natural frequency is manifested by a peak pressure, as expected.


Figure 2.
To explain why the average velocity does not show a resonance at some natural frequencies, we must examine Eqs. (5) and (7) more carefully. The former indicates that the $j$ th pressure term will be singular if at some frequency, $J_{1}\left(\mu_{0 . j} a\right)=0$, which encompasses the alternative that the factor $\mu_{0 . j}=0$. This is the attribute of any natural frequency. The situation for $V_{\mathrm{av}}$ in Eq. (7) is different, because a singularity only
occurs if the frequency is such that one of the $\mu_{0 . j}$ terms in the denominator is zero. The preceding tabulation indicates $\mu_{0 . j}=0$ is associated with the natural frequencies at 170,510 , and 850 Hz . This merely confirms that the graphs are consistent with the analysis, but not why the pressure and average velocity behave differently. The answer lies in Eq. (6) for the axial particle velocity at any location. The denominator of that quantity contains $J_{1}\left(\mu_{0, j} a\right)$, so it is singular at any natural frequency. However, the numerator indicates that the velocity distribution at $x=0$ is proportional to $J_{0}\left(\mu_{0, j} R\right)$. The average value of this function over the area is proportional to $J_{1}\left(\mu_{0 . j} a\right)$, which cancels the singular denominator term. Recall that $\mu_{0, j}=0$ is the property of a plane wave mode, in which there is no dependence on $R$. Hence, we have established that if any mode is not constant over a cross section, its contribution to the average velocity will be finite at any frequency, even resonant ones. The average velocity will show a resonance only if the resonant mode is one in which there is no variation in the $R$ direction.

To close this example, let us consider whether this device is useful as a lowfrequency loudspeaker. Flatness of the frequency response is a desirable attribute. Filling the cavity with air leads to many resonances in the frequency band, but the average velocity in the water-filled design is essentially flat out to 500 Hz . This might be a desirable attribute, but another feature of the response is not. In the low-frequency range, where the response is flat, the value of $V_{\mathrm{av}}$ is nearly equal to $V_{a}$. Thus, no advantage is gained by driving the cylindrical surface relative to using a piston that matches the cylinder's cross section. Resonances and amplification factors are two aspects that must be considered in the design of any transducer.

### 10.3 Analysis Using Natural Cavity Modes

The forced cavity modes depend on the frequency of excitation and what surface is forced to vibrate. In contrast, the natural cavity modes that are the present topic depend on the size and shape of the enclosed region, as well as the nature of the boundaries, but they do not depend on the way in which the field is generated. These modes are properties of the system.

They must be determined prior to analysis of the response to an excitation. If the region has a shape that fits one of the standard coordinate systems, there will be wavenumbers describing the spatial dependence in each coordinate direction. In that case, the mode will be characterized by a subscript whose value marks the set of associated wavenumbers. For example, the modes of a rectangular room would be denoted as $\Psi_{j, m, n}(x, y, z)$. However, many situations feature an enclosed region whose shape is irregular They do not lend themselves to decomposition into identifiable wavenumbers. Such systems are better described by a single subscript denoting the hierarchy of modes. We will begin with an overview, which is more easily presented if we employ the latter representation. Hence, the cavity modes in the general development will be denoted as $\Psi_{n}(\bar{x})$.

Before we embark, it is necessary to limit the scope of our investigations to situations where the walls fit some combination of the rigid or pressure-release models.

This restriction is imposed because the formulation required to treat any other wall condition is quite complicated, as is evident from a fairly recent analysis. ${ }^{5}$ Even under this constraint, a variety of configurations are possible. The development that follows begins with the general aspects of cavity modes that do not depend on the shape and properties of the enclosure. These are the governing equations, orthogonality, and the usage of a modal series to determine the pressure. The general development will be followed by specific analyses of standard shapes. Another limitation is that only the response generated by sources situated within the fluid will be considered. Recall that this is the situation for which a formulation using natural cavity modes is best suited, whereas the field generated by an imposed wall vibration is best described in terms of forced cavity modes. The common practice is to omit "natural" from the terminology. We shall do so because each formulation is used for a different class of excitation. Consequently, there is no context where both natural and forced cavity modes will be encountered.

### 10.3.1 Equations Governing Cavity Modes

By definition, a cavity mode $\Psi_{n}$ is a pressure field that can exist within the fluid domain $\mathcal{V}$, even though there is no external excitation. Such a function must be a solution of the Helmholtz equation. In accord with the restrictions stated above, we subdivide the boundary into a rigid portion $\mathcal{S}_{V}$ on which the normal velocity is zero and a pressure-release portion $\mathcal{S}_{P}$ on which the pressure is zero. The pressure field in a cavity mode is $p(\bar{x}, t)=\operatorname{Re}\left(\Psi_{n} \exp (i \omega t)\right)$, and Euler's equation describes the particle velocity. Hence, the equations governing a cavity mode for this set of surfaces are

$$
\begin{gather*}
\nabla^{2} \Psi_{n}+\left(k_{n}\right)^{2} \Psi_{n}=0, \quad \bar{x} \in \mathcal{V} \\
\bar{n}\left(\bar{x}_{s}\right) \cdot \nabla \Psi_{n}=0 \text { if } \bar{x}_{s} \in \mathcal{S}_{V}  \tag{10.3.1}\\
\Psi_{n}=0 \text { if } \bar{x}_{s} \in \mathcal{S}_{P}
\end{gather*}
$$

If we assign an arbitrary value to $k_{n}$, the only possible solution of this homogeneous set of equations is $\Psi_{n}=0$ everywhere. Thus, we must identify the values of $k_{n}$ for which there is a nontrivial solution. This is a statement of an eigenvalue problem. The eigenvalues $k_{n}$ give the natural frequencies as $\omega_{n}=k_{n} c$, and the eigenfunction corresponding to each $k_{n}$ is a cavity mode.

It should be evident that it will be quite difficult, or even impossible, to solve for the eigenvalues $k_{n}$ unless $\mathcal{V}$ has a shape that fits one of the standard coordinate systems. Thus, we shall pursue analytical solutions only for rectangular, spherical, and cylindrical cavities. Any combination of rigid and pressure-release subregions of $\mathcal{S}$ is admissible from the viewpoint of general properties. However, we will not be able to solve analytically for the mode functions if the condition on a constant coordinate surface changes. For example, we can use analytical methods if entire

[^37]walls of a box are pressure-release or rigid, but we cannot use analytical methods if any wall of that box is partially rigid and partially pressure-release. Such situations require computational or approximate methods.

The existence and nature of a zero frequency mode has general importance. If the frequency is zero, the Helmholtz equation reduces to Laplace's equation, $\nabla^{2} P=0$. Because $\nabla^{2} P \equiv \nabla \cdot \nabla P$, it must be that $\nabla P$ is a constant vector. If the entire boundary is rigid, then $\bar{n} \cdot \nabla P$ must be zero at the boundary, which is satisfied only if the constant value of $\nabla P$ is zero. In other words, a constant value of $P$ constitutes a nontrivial solution at zero frequency if a fluid is fully enclosed by a rigid boundary. The contrary case is that in which any portion of the boundary is pressure-release. A field whose gradient is constant with a zero value at any location must be zero everywhere. In other words,

If a region of fluid is fully enclosed by a rigid boundary, then there exists a zero frequency cavity mode in which the pressure is constant. If any portion of the boundary is pressure-release, a zero frequency mode does not exist.

### 10.3.2 Orthogonality

Orthogonality of the cavity modes greatly simplifies determination of the response to source excitation. To derive this property, we consider a pair of modes $j$ and $n$, whose natural frequencies are $c k_{j}$ and $c k_{n}$. The first step entails multiplying the Helmholtz equation for mode $n$ by mode $j$ and then integrating over $\mathcal{V}$, which gives

$$
\begin{equation*}
\iiint_{\mathcal{V}} \Psi_{j} \nabla^{2} \Psi_{n} d \mathcal{V}+k_{n}^{2} \iint_{\mathcal{V}} \int_{j} \Psi_{n} d \mathcal{V}=0 \tag{10.3.2}
\end{equation*}
$$

In order to obtain a form in which $\Psi_{j}$ and $\Psi_{n}$ appear in the same manner, we exploit the definition that $\nabla^{2} \Psi \equiv \nabla \cdot \nabla \Psi$ to convert the first integrand,

$$
\begin{equation*}
\iint_{\mathcal{V}} \int\left[\nabla \cdot\left(\Psi_{j} \nabla \Psi_{n}\right)-\nabla \Psi_{j} \cdot \nabla \Psi_{n}\right] d \mathcal{V}+k_{n}^{2} \iiint_{\mathcal{V}} \Psi_{j} \Psi_{n} d \mathcal{V}=0 \tag{10.3.3}
\end{equation*}
$$

Application of the divergence theorem to the first term leads to

$$
\begin{equation*}
-\iint_{\mathcal{S}} \Psi_{j} \bar{n}\left(\bar{x}_{S}\right) \cdot \nabla \Psi_{n} d \mathcal{S}-\iiint_{\mathcal{V}} \nabla \Psi_{j} \cdot \nabla \Psi_{n} d \mathcal{V}+k_{n}^{2} \iiint_{\mathcal{V}} \Psi_{j} \Psi_{n} d \mathcal{V}=0 \tag{10.3.4}
\end{equation*}
$$

where the negative sign preceding the surface integral results from the definition of $\bar{n}$ as the normal oriented into the cavity.

We have limited our consideration to systems in which the wall consists of a rigid portion $\mathcal{S}_{V}$ and a pressure-release portion $\mathcal{S}_{P}$. On $\mathcal{S}_{V}$, we have $\bar{n}\left(\bar{x}_{s}\right) \cdot \nabla \Psi_{n}=0$ and $\Psi_{n}\left(\bar{x}_{s}\right)=0$ on $\mathcal{S}_{P}$. In either case, one of the factors in the surface integral is zero, so it must be that

$$
\begin{equation*}
k_{n}^{2} \iiint_{\mathcal{V}} \Psi_{j} \Psi_{n} d \mathcal{V}-\iiint_{\mathcal{V}} \nabla \Psi_{j} \cdot \nabla \Psi_{n} d \mathcal{V}=0 \tag{10.3.5}
\end{equation*}
$$

If we had begun by multiplying the Helmholtz equation for mode $j$ by mode $n$, we would have obtained

$$
\begin{equation*}
k_{j}^{2} \iiint_{\mathcal{V}} \Psi_{j} \Psi_{n} d \mathcal{V}=\iiint_{\mathcal{V}} \nabla \Psi_{n} \cdot \nabla \Psi_{j} d \mathcal{V} \tag{10.3.6}
\end{equation*}
$$

The difference of these two relations is

$$
\begin{equation*}
\left(k_{n}^{2}-k_{j}^{2}\right) \iiint_{\mathcal{V}} \Psi_{j} \Psi_{n} d \mathcal{V}=0 \tag{10.3.7}
\end{equation*}
$$

This relation is reminiscent of the one for the transverse mode function of a forced cavity mode. If modes $j$ and $n$ are different, then the factor multiplying the integral is not zero, so the integral must vanish. In other words, two different modes are orthogonal over the space $\mathcal{V}$. Equation (10.3.7) also applies if $\Psi_{n}$ and $\Psi_{j}$ are the same function. In that case, the factor multiplying the integral is zero, so the condition is met regardless of the value of the integral. Indeed, we know that the integration will yield a positive value because the mode functions are real, so $\left(\Psi_{n}\right)^{2}>0$ everywhere. This positive value is used to normalize the cavity modes. Specifically, a cavity mode $\Psi_{n}$ may be multiplied by any constant $B_{n}$ without altering the fact that they are solutions of the basic governing equations. The arbitrariness of $B_{n}$ is addressed by setting it such that the integral equals the volume $\mathcal{V}$ of the domain. Doing so yields mode functions that are dimensionless. Thus, we have the first orthogonality condition,

$$
\begin{equation*}
\iiint_{\mathcal{V}} \Psi_{j} \Psi_{n} d \mathcal{V}=\mathcal{V} \delta_{j, m} \tag{10.3.8}
\end{equation*}
$$

(For two-dimensional domains, $\mathcal{V}$ would be the area in the plane on which the pressure varies.) When we substitute the first orthogonality condition into Eq. (10.3.5), we obtain the second orthogonality condition,

$$
\begin{equation*}
\iiint_{\mathcal{V}} \nabla \Psi_{j} \cdot \nabla \Psi_{n} d \mathcal{V}=\mathcal{V}\left(k_{n}\right)^{2} \delta_{j, m} \tag{10.3.9}
\end{equation*}
$$

An exception to this development arises if two modes have the same natural frequency. This situation is common for rectangular-shaped enclosures, as we will see in the upcoming Example 10.7. Furthermore, the repeated natural frequencies always arise for the nonaxisymmetric modes of a cylindrical cavity because of the arbitrari-
ness of how $\theta$ is defined. However, in both geometries, the derived cavity modes will be orthogonal, even though $k_{j}=k_{n}$. This property stems from the differing spatial patterns, with one being obtainable from the other by rotating the coordinate system by $90^{\circ}$ from the other.

### 10.3.3 Analysis of the Pressure Field

We may now proceed to formulate a modal analysis of the pressure induced by sources. One of the assets of the cavity mode approach is that it can be used to derive directly a time-domain response. We shall develop this type of solution and then use the result to obtain a description of the frequency-domain response.

We begin with a series of cavity modes in which the modal coefficients are allowed to be arbitrary functions of time,

$$
\begin{equation*}
p(\bar{x}, t)=\sum_{n=1}^{\infty} \Psi_{n}(\bar{x}) p_{n}(t) \tag{10.3.10}
\end{equation*}
$$

An arbitrary collection of sources are situated within the cavity. The volume velocity of each source is $Q_{m}(t)$, and the locations are $\bar{x}_{m}$. The field in the time domain for a single source is governed by the inhomogeneous wave equation, Eq. (6.4.23), with $C=-\rho_{0} \dot{Q}_{m}$. Thus, the field equation for the pressure is

$$
\begin{equation*}
\nabla^{2} p-\frac{1}{c^{2}} \frac{\partial^{2} p}{\partial t^{2}}=-\rho_{0} \sum_{m} \dot{Q}_{m}(t) \delta\left(\bar{x}-\bar{x}_{m}\right) \tag{10.3.11}
\end{equation*}
$$

Differentiation of the modal series with respect to $t$ operates only on the modal coefficients, and spatial derivatives operate only on the cavity modes. Thus, substitution of the series into the preceding wave equation leads to

$$
\begin{equation*}
\sum_{n=1}^{\infty}\left[c^{2} \nabla^{2} \Psi_{n} p_{n}-\Psi_{n} \ddot{p}_{n}\right]=-\rho_{0} c^{2} \sum_{m} \dot{Q}_{m}(t) \delta\left(\bar{x}-\bar{x}_{m}\right) \tag{10.3.12}
\end{equation*}
$$

The procedure mirrors that which led to Eq. (9.2.52). The preceding is multiplied by a specific $\Psi_{j}$ and integrated over $\mathcal{V}$. Because $\Psi_{j} \nabla^{2} \Psi_{n} \equiv \nabla \cdot\left(\Psi_{j} \nabla \Psi_{n}\right)-\nabla \Psi_{j} \cdot \nabla \Psi_{n}$, the divergence theorem gives

$$
\begin{equation*}
\iiint_{\mathcal{V}} \Psi_{j} \nabla^{2} \Psi_{n n} d \mathcal{V}=\iint_{\mathcal{S}} \Psi_{j} \bar{n} \cdot \nabla \Psi_{n} d \mathcal{S}-\iiint_{\mathcal{V}} \nabla \Psi_{j} \cdot \nabla \Psi_{n} d \mathcal{V} \tag{10.3.13}
\end{equation*}
$$

The mode functions satisfy rigid or pressure-release boundary conditions, as prescribed in Eq. (10.3.1), so the surface integral vanishes. Hence, we have found that

$$
\begin{array}{r}
\sum_{n=1}^{\infty} c^{2} p_{n} \iiint_{\mathcal{V}} \nabla \Psi_{j} \cdot \nabla \Psi_{n} d \mathcal{V}+\sum_{n=1}^{\infty} \ddot{p}_{n} \iiint_{\mathcal{V}} \Psi_{j} \Psi_{n} d \mathcal{V} \\
=\iiint_{\mathcal{V}} \Psi_{j}\left[\rho_{0} c^{2} \sum_{m} \dot{Q}_{m}(t) \delta\left(\bar{x}-\bar{x}_{m}\right)\right] d \mathcal{V} \tag{10.3.14}
\end{array}
$$

The last step is to invoke both orthogonality properties, Eqs. (10.3.8) and (10.3.9), which serve to filter out of both summations all terms for which $n \neq j$. The remaining terms constitute an ordinary differential equation for the $j$ th modal coefficient,

$$
\begin{equation*}
\ddot{p}_{j}+\left(\omega_{j}\right)^{2} p_{j}=S_{j}(t), \quad j=1,2, . . \tag{10.3.15}
\end{equation*}
$$

The $S_{j}$ excitations are time-dependent modal source strengths. They map the spatial distribution of sources into the modal space, according to

$$
\begin{equation*}
S_{j}=\frac{\rho_{0} c^{2}}{\mathcal{V}} \iiint_{\mathcal{V}} \Psi_{j} \sum_{m} \dot{Q}_{m}(t) \delta\left(\bar{x}-\bar{x}_{m}\right) d \mathcal{V}=\frac{\rho_{0} c^{2}}{\mathcal{V}} \sum_{m} \dot{Q}_{m}(t) \Psi_{j}\left(\bar{x}_{m}\right) \tag{10.3.16}
\end{equation*}
$$

Equation (10.3.15) has the same form as the differential equation of motion for a simple spring-mass system. Its solution is quite accessible. If there is a single source whose time signature is a unit impulse function, that is, $\dot{Q}_{1}=\delta\left(t-t_{0}\right)$, then the modal coefficients are the impulse response,

$$
\begin{equation*}
p_{j}=\frac{\rho_{0} c^{2}}{\mathcal{V}} \frac{\Psi_{j}\left(\bar{x}_{1}\right)}{\omega_{j}} \sin \left(\omega_{j}\left(t-t_{0}\right)\right) h\left(t-t_{0}\right) \tag{10.3.17}
\end{equation*}
$$

By definition, the time-domain Green's function is the pressure due to an impulsive point source. Thus, we have established that for a cavity it is

$$
\begin{equation*}
G(\bar{x}, t)=\frac{\rho_{0} c^{2}}{\mathcal{V}} \sum_{j=1}^{\infty} \frac{\Psi_{j}\left(\bar{x}_{1}\right) \Psi_{j}(\bar{x})}{\omega_{j}} \sin \left(\omega_{j}\left(t-t_{0}\right)\right) h\left(t-t_{0}\right) \tag{10.3.18}
\end{equation*}
$$

Depending on the nature of the sources' time signatures, the modal differential equations might be solvable by standard methods. If not, they can be solved numerically. One case where an analytical solution is possible is that of steady-state field generated by time harmonic sources. Such an analysis yields the frequency-domain solution. The sources are set to $\dot{Q}_{m}(x, t)=\delta\left(\bar{x}-\bar{x}_{m}\right) \operatorname{Re}\left[i \omega \hat{Q}_{m} \exp (i \omega t)\right]$. The frequency-domain representation of the pressure coefficients is $p_{j}=\operatorname{Re}\left(X_{j} e^{i \omega t}\right)$. The modal source strengths in Eq. (10.3.16) become

$$
\begin{equation*}
S_{n}=\sum_{m} \Psi_{n}\left(\bar{x}_{m}\right) \operatorname{Re}\left(\frac{i \omega \rho_{0} c^{2} \hat{Q}_{m}}{\mathcal{V}} e^{i \omega t}\right) \tag{10.3.19}
\end{equation*}
$$

Synthesis of the modal series, Eq. (10.3.10), using the particular solution for the $P_{j}$ coefficients yields

$$
\begin{equation*}
p(\bar{x}, t)=\rho_{0} c^{2} \sum_{n=1}^{\infty} \sum_{m} \frac{\Psi_{n}\left(\bar{x}_{m}\right) \Psi_{n}(\bar{x})}{\omega_{n}^{2}-\omega^{2}} \operatorname{Re}\left(\frac{i \omega \hat{Q}_{m}}{\mathcal{V}} e^{i \omega t}\right) \tag{10.3.20}
\end{equation*}
$$

In the special case of a single source whose strength is a unit value, $\hat{Q}_{1}=1$, the preceding gives the frequency-domain Green's function of the cavity. The similarity of this result and Eq. (10.1.16), which was derived heuristically, is obvious.

The general development is applicable to the cavity modes of any configuration, regardless of the shape of the enclosed region and the nature of the walls. However, Eq. (10.3.20) presumes that the cavity modes have been determined. The reality is that only rectangular and circular shapes may be analyzed without considerable effort.

### 10.3.4 Rectangular Cavity

To study the field within a closed box, we place the origin of $x y z$ at one corner, with the axes aligned parallel to the edges. Then, the sides of the cavity coincide with the planes $x=0, x=L, y=0, y=H, z=0$, and $z=W$. The pressure in the region between two parallel walls was the subject of Sect. 9.2. It was shown there that any combination of rigid and pressure-release surfaces is represented by either a sine or cosine function of the coordinate that is normal to the surfaces. That is the present situation, but there are now three sets of parallel walls. Consequently, there are three sets of harmonic functions: $\Psi_{x}(x)$ for the walls at $x=0$ and $x=L, \Psi_{y}(y)$ for the walls at $y=0$ and $y=H$, and $\Psi_{z}(z)$ for the walls at $z=0$ and $z=W$. The wavenumbers are denoted as $\kappa, \mu$, and $\nu$ in the $x, y$, and $z$ directions, respectively. A product of these harmonic functions will satisfy the Helmholtz equation if the wavenumbers are set properly. Furthermore, because the differential equation and boundary conditions are homogeneous, any solution may be multiplied by a constant $B$. It follows from these considerations that a mode function is described by

$$
\begin{equation*}
\Psi(x, y, z)=B \Psi_{x}(x) \Psi_{y}(y) \Psi_{z}(z) \tag{10.3.21a}
\end{equation*}
$$

where the factors are

$$
\begin{align*}
& \Psi_{x}=c_{x} \cos (\kappa x)+s_{x} \sin (\kappa x) \\
& \Psi_{y}=c_{y} \cos (\mu y)+s_{y} \sin (\mu y) \\
& \Psi_{z}=c_{z} \cos (\nu z)+s_{z} \sin (\nu z) \tag{10.3.21b}
\end{align*}
$$

Either a $c$ or $s$ coefficient will be zero, depending on the nature of the side that contains the origin. The nonzero coefficient may be set to one because $B$ serves to collect all arbitrary factors. For example, if the surface $y=0$ is rigid, then it must be that $\partial P / \partial y=0$ at $y=0$, which is met by setting $s_{y}=0$ and $c_{y}=1$. Alternatively, if $y=0$ is a pressure-release side, so that $P=0$ at $y=0$, then $c_{y}=0$ and $s_{y}=1$.

The wavenumber in each direction is found by satisfying the boundary conditions at the far walls, $x=L, y=H$, and $z=W$. In particular, the condition that there be a nontrivial solution for each direction will require that the argument of the respective sinusoidal function be an integer multiple of $\pi / 2$. For example, suppose $P=0$ on $z=0$, so that $\Psi_{z}=\sin (\nu z)$. Then, the rigid condition, $\partial P / \partial z=0$ on $z=W$, requires that $\cos (\nu W)=0$, which is satisfied if $\nu W$ is an odd multiple of $\pi / 2$. The other possibility, $P=0$ at $z=W$, requires $\sin (\nu W)=0$. This is met if $\nu W$ is an even multiple of $\pi$, but $\nu$ cannot be zero. This is so because it would lead to the trivial solution $\Psi=0$.

Determination of the possible values of the wavenumbers allows us to evaluate the natural frequencies. By definition, Eq.(10.3.21) is a solution of the Helmholtz equation at a natural frequency. Let $j, m, n$ be the sequence number for $k, \mu$, and $v$, respectively, so that the corresponding natural frequency is denoted as $\omega_{j, m, m}=$ $c k_{j, m, n}$. Then, it must be that

$$
\begin{equation*}
\omega_{j, m, m}=c\left(\kappa_{j}^{2}+\mu_{m}^{2}+\nu_{n}^{2}\right)^{1 / 2} \tag{10.3.22}
\end{equation*}
$$

This leaves only the scaling coefficient $B_{j, m, n}$ to determine. These values are obtained by normalizing the mode functions to satisfy Eq.(10.3.8) when the two modes are the same. The mode function described in Eq. (10.3.21) is products of the functions of each coordinate, and the boundaries consist of constant values of each coordinate. Hence, Eq. (10.3.8) leads to

$$
\begin{equation*}
B_{j, m, n}=\left(\frac{L H W}{\int_{0}^{W} \Psi_{x}(x)^{2} d x \int_{0}^{H} \Psi_{y}(y)^{2} d y \int_{0}^{L} \Psi_{z}(z)^{2} d z}\right)^{1 / 2} \tag{10.3.23}
\end{equation*}
$$

A closed form expression for the coefficient may be found, but it is easier to do so after the functions are specified.

EXAMPLE 10.7 A column of water is used to test vertical propagation in a channel. The height is 4 m , and the dimensions of the horizontal cross section are $0.6 \mathrm{~m} \times 1.2 \mathrm{~m}$. The bottom and sides of the tank are rigid. Determine the natural frequencies of this tank up to 4 kHz .

## Significance

The evaluation of the frequencies is straightforward. What makes this example interesting is what the results tell us about the general properties of natural frequencies.

## Solution

We shall align the $y$-axis vertically and set the $x$-axis horizontally along the short side. This corresponds to $L=0.6 \mathrm{~m}, H=4 \mathrm{~m}$, and $W=1.2 \mathrm{~m}$. The surface $y=H$ is the interface between the water and air, so it is well approximated as pressurerelease for waves in the water. The vertical sides and bottom are rigid, so it is required that the normal derivatives on them be zero. Satisfaction of this condition at $x=0$, $y=0$, and $z=0$ requires that

$$
\begin{equation*}
\Psi_{x}=\cos (\kappa x), \quad \Psi_{y}=\cos (\mu y), \quad \Psi_{z}=\cos (\nu z) \tag{1}
\end{equation*}
$$

The other vertical sides are $x=L$ and $z=W$. Setting the normal derivative on each to zero yields

$$
\left.\frac{\partial \Psi_{x}}{\partial x}\right|_{x=L}=\sin (\kappa L)=0 \text { and }\left.\frac{\partial \Psi_{z}}{\partial z}\right|_{z=W}=\sin (\nu W)=0
$$

These conditions are met if

$$
\begin{equation*}
\kappa_{j}=\frac{j \pi}{L}, \quad j=0,1, \ldots ; \quad \nu_{n}=\frac{n \pi}{W}, \quad n=0,1, \ldots \tag{2a}
\end{equation*}
$$

Note that $\kappa_{0}=0$ and $\nu_{0}=0$. Both correspond to the plane wave modes of twodimensional waveguide whose walls are rigid. The horizontal surface at $y=H$ is pressure-release, so setting $\Psi_{y}=0$ there leads to

$$
\begin{equation*}
\mu_{m}=\frac{(2 m-1) \pi}{2 H}, \quad m=1,2, \ldots \tag{2b}
\end{equation*}
$$

The natural frequency corresponding to these wavenumbers is

$$
\begin{equation*}
\omega_{j, m, n}=\pi c\left[\left(\frac{j}{L}\right)^{2}+\left(\frac{2 m-1}{2 H}\right)^{2}+\left(\frac{n}{W}\right)^{2}\right]^{1 / 2} \tag{3}
\end{equation*}
$$

The mode functions, which are not needed in the present context, are

$$
\begin{equation*}
\Psi_{j, m, n}=B_{j, m, n} \cos \left(\frac{j \pi x}{L}\right) \cos \left(\frac{(2 m-1) \pi y}{2 H}\right) \cos \left(\frac{n \pi z}{W}\right) \tag{4}
\end{equation*}
$$

It is specified that we should limit our attention to the range up to 4 kHz . To determine the range of indices to consider, we observe that the natural frequency increases monotonically as either $j, m$, or $n$ is increased with the other two held fixed.

For a fixed frequency, the largest value of one index corresponds to the smallest value of the other two. The range of indices for the computation is $0 \leq j \leq J, 1 \leq m \leq M$, $0 \leq n \leq N$. Thus, we set the range such that $J, M$, and $N$ are the smallest values for which $\omega_{J, 1,0}>\omega_{\max }, \quad \omega_{0, M, 0}>\omega_{\max }$, and $\omega_{0,1, N}>\omega_{\max }$. Setting $\omega_{\max }=8000 \pi$ $\mathrm{rad} / \mathrm{s}$ leads to $J=4, M=23$, and $N=7$.

A four column list of $j, m, n$, and $\omega_{j, m, n}$ for $j=0, . ., 3, m=1, \ldots, 22$, and $n=0, . ., 6$ could be sorted by saving them to a spreadsheet. A graphical display is a better way to see the qualitative behavior. Each curve in Fig. 1 describes $\omega_{j, m, n}$ for a fixed $j$ and $n$ as $m$ is increased from one. The lowest frequency, which corresponds to $j=0, m=1$, and $n=0$, is 92.5 Hz . This is the cavity's fundamental natural frequency.


Figure 1.
A few combinations of $j$ and $n$ correspond to overlapping curves, which means the natural frequencies are equal regardless of $m$. The lowest curve for which this condition occurs is $(0, m, 2)$ and $(1, m, 0)$. Other combinations are $(0, m, 4)$ and $(2, m, 0),(1, m, 4),,(2, m, 2$,$) , and so on. This overlap is a consequence of the fact$ that $W=2 L$. Thus, if a natural frequency corresponds to indices $j_{1}, m$, and $n_{1}$, the same value will occur for $j_{2}=n_{1} / 2$ and $n_{2}=2 j_{1}$, with the same $m$. This is not to suggest that natural frequencies overlap only if $j_{2} n_{2}=j_{1} n_{1}$. Other combinations are possible. This is the case for the $(0, m, 5)$ and $(2, m, 3)$ modes. Furthermore, even if a frequency curve does do not overlap others, it is possible that a value on that curve will equal a value on a different curve at a different $m$. An interesting manifestation of that possibility is a triple value, $\omega_{1,8,3} /(2 \pi)=\omega_{1,13,0} /(2 \pi)=\omega_{0,13,2} /(2 \pi)=2621 \mathrm{~Hz}$.

In general, different modes will have the same natural frequency in the lowfrequency range if a ratio of cavity dimensions is a small rational fraction. For example, in the present system the ratio of lengths in the $x$ and $z$ directions is $W / L=2$, so the occurrence of equal frequencies with different $(j, n)$ pairs is rather common. In contrast, $H / L=20 / 3$, so equal natural frequencies with different ( $j, m$ )
are less common. The same is true for natural frequencies associated with different $(m, n)$ pairs because $H / W=10 / 3$.

It was stated in the general discussion that two cavity modes will be orthogonal even if they have the same natural frequency. Let us test that assertion. The orthogonality condition is the volume integral of the product of two functions. In view of the fact that the modes in Eq. (4) are products of $x, y$, and $z$ functions, the orthogonality integral factorizes to integrals over the range of $x, y$ and $z$. Each has the same general form,

$$
\int_{0}^{\ell} \cos \left(\beta_{r} \eta\right) \cos \left(\beta_{s} \eta\right) d \eta \equiv \frac{1}{2}\left[\frac{\sin \left(\left(\beta_{r}-\beta_{s}\right) \ell\right)}{\beta_{r}-\beta_{s}}+\frac{\sin \left(\left(\beta_{r}+\beta_{s}\right) \ell\right)}{\beta_{r}+\beta_{s}}\right]
$$

For the $x$ functions, $\beta_{r} \ell$ represents $\kappa_{j} L=j \pi$, while $\beta_{r} \ell$ represents $\nu_{n} W=n \pi$ for the $z$ functions. In both cases, the argument of the above sine functions is a multiple of $\pi$, and therefore zero. For the $y$ functions, $\beta_{r} \ell$ represents $\mu_{m} H=(m-1 / 2) \pi$, so that $\left(\beta_{r} \pm \beta_{s}\right) \ell$ also is a multiple of $\pi$. It follows that the orthogonality integral vanishes for any combination of two different mode functions, regardless of what their natural frequencies are.

The fact that two or more modes share a natural frequency does have some interesting implications. Suppose a field is established at the natural frequency belonging to two modes. A linear combination of these modes also satisfies the governing equations. For example, in the case of a pair identified in the preceding paragraph, $\Psi=b_{1} \Psi_{0,1,2}+b_{2} \Psi_{1,1,0}$ satisfies the Helmholtz at $k=\omega_{0,1,2} / c=\omega_{1,1,0} / c$ for any constants $b_{1}$ and $b_{2}$. We could consider the combined function to be a mode, but shall not do so because it is not independent of the ones we derived. The values of $b_{1}$ and $b_{2}$ that occur depend on how the field is induced. This phenomenon was first encountered in the vibration of rectangular membranes and plates. Chaladni ${ }^{6}$ placed sand on a vibrating square plate. This made it possible to visualize the vibration pattern because the sand migrated toward the nodal lines, at which the plate is quiet. The excitation consisted of a violin bow that was stroked across an edge. Chaladni showed that nodal patterns were not unique. He also showed that they were very difficult to reproduce. This is a consequence of the fact that the factors $b_{1}$ and $b_{2}$ for the vibration were highly sensitive to small changes in the location where the plate was stroked. In recognition of his contributions, nodal lines are commonly known as Chaladni lines.

Even if natural frequencies are not exactly equal, the likelihood that two or more natural frequencies are very close increases drastically as the frequency increases. This attribute may be recognized by drawing a horizontal line through any natural frequency in the graph. Above $\omega_{0,1,1}$, this line will intersect two or more curves. As the frequency increases, the number of curves such as line intersects will increase. Some of those intersections will be close to integer values of $m$, which mark the natural frequencies. We say that the modal density increases with increasing natural

[^38]frequency. This phenomenon is a general attribute of two- and three-dimensional systems that conserve energy. It can be quite problematic for experimental methods that attempt to identify modal properties from measured data.

EXAMPLE 10.8 Consider a hard-walled chamber whose dimensions are $4 \times 3 \times 6 \mathrm{~m}$. A line source is situated along one of the 6 m edges. The source is impulsive, with a volume velocity per unit length that changes discontinuously at $t=0$ from zero to a constant value $q_{0}$. This corresponds to a mass acceleration per unit length that is $\rho_{0} q_{0} \delta(t)$. Determine the signal heard in the middle of the chamber. The fluid is air.

## Significance

The solution is a direct application of the modal series formalism. The results will lead to some useful insights regarding our prior study of point sources, and it will highlight some of the difficulties one might encounter in the application of modal series representations.

## Solution

The $x y z$ coordinate system in Fig. 1 has the $z$-axis coincident with the line source and the $x$-axis coincident with a 4 m edge. The line source has the same properties along its length, and the hard-walls at $z=0$ and $z=6 \mathrm{~m}$ allow for the existence of a pressure field that does not vary in that direction. Thus, the pressure field is $p=\operatorname{Re}(P(x, y) \exp (i \omega t))$.


Figure 1.

Despite independence of the $z$ coordinate, it is best to treat the field as threedimensional, rather than considering it to be two-dimensional. The boundaries in the $x$ and $y$ directions also are rigid, so the modes that contribute to the pressure field are

$$
\begin{equation*}
\Psi_{j, n, 0}=B_{j, n, 0} \cos \left(\frac{j \pi x}{L}\right) \cos \left(\frac{n \pi y}{H}\right), \quad j, n=0,1, \ldots \tag{1}
\end{equation*}
$$

where $L=4 \mathrm{~m}$ and $H=3 \mathrm{~m}$. The corresponding natural frequencies are

$$
\begin{equation*}
\omega_{j, n, 0}=c\left[\left(\frac{j \pi}{L}\right)^{2}+\left(\frac{n \pi}{H}\right)^{2}\right]^{1 / 2} \tag{2}
\end{equation*}
$$

Normalizing the mode functions to the volume $\mathcal{V}=L H W$, in accord with Eq. (10.3.8), leads to

$$
\begin{equation*}
B_{j, n, 0}=\frac{2}{\left(1+\delta_{0, j}\right)^{1 / 2}\left(1+\delta_{0, n}\right)^{1 / 2}} \tag{3}
\end{equation*}
$$

Because the source is situated in the corner, its mass flows into one quadrant, rather than into free space. To account for this effect, we multiply the source strength by a factor of four. Another consideration is that Eq. (10.3.16) describes the modal source strength $S_{j}$ for a discrete set of sources. The line source is a continuous distribution of differential mass accelerations $4 \rho_{0} q_{0} \delta(t) d z$, so a summation of sources is replaced by an integral over $z$. Hence, the first form of Eq. (10.3.16) in this case gives

$$
S_{j, n, 0}=\frac{\rho_{0} c^{2}}{L H W} \int_{0}^{L} \int_{0}^{H} \int_{0}^{W}\left\{\int_{0}^{W} 4 \rho_{0} q_{0} \delta(t) d \xi\right\} \delta(x) \delta(y) \Psi_{j, n, 0} d z d y d x
$$

Equation (1) indicates that all modes at the origin are $\Psi_{j, n, 0}=B_{j, n, 0}$. Thus, the modal source strengths are

$$
\begin{equation*}
S_{j, n, 0}=\frac{4 \rho_{0} c^{2} q_{0}}{L H} B_{j, n, 0} \delta(t) \tag{4}
\end{equation*}
$$

The modal differential equations, Eq.(10.3.15), corresponding to these coefficients are

$$
\begin{equation*}
\ddot{p}_{j, n, 0}+\left(\omega_{j, n, 0}\right)^{2} p_{j, n, 0}=\frac{4 \rho_{0} c^{2} q_{0}}{L H} B_{j, n, 0} \delta(t) \tag{5}
\end{equation*}
$$

Quiescent initial conditions are $p_{j, n, 0}=0$ and $\dot{p}_{j, n, 0}=0$ for any $t$ preceding activation of the impulse. After the impulse, the excitation is gone, so only the homogeneous solution remains. The initial conditions for the ensuing response are obtained by considering the infinitesimal interval from $t=0^{-}$before the impulse to $t=0^{+}$ afterward. The pressure cannot change suddenly in that interval, because doing so would lead to an infinite particle acceleration. Thus, it must be that $p_{j, n}\left(t=0^{+}\right)=0$. To determine $\dot{p}_{j, n}\left(t=0^{+}\right)$, we integrate Eq. (5) over $0^{-}<t<0^{+}$. The integral of $p_{n, n}$ over this infinitesimal interval is zero, and the integral of $\ddot{p}_{j, n}$ is $\dot{p}_{j, n}\left(t=0^{+}\right)-\dot{p}_{j, n}\left(t=0^{-}\right)$. Thus, we find that

$$
\dot{p}_{j, n, 0}\left(t=0^{+}\right)=\frac{4 \rho_{0} c^{2} q_{0}}{L H} B_{j, n, 0}
$$

The solution of Eq. (5) fitting the initial conditions is the one-degree-of-freedom impulse response,

$$
\begin{align*}
P_{j, n, 0} & =\frac{4 \rho_{0} c^{2} q_{0}}{L H} B_{j, n, 0} g_{j, n, 0}(t) \\
g_{j, n, 0} & =\frac{\sin \left(\omega_{j, n, 0} t\right)}{\omega_{j, n, 0}} h(t) \text { if } j>0 \text { or } n>0  \tag{6}\\
g_{0,0,0} & =\operatorname{th}(t)
\end{align*}
$$

where the step function serves to remind us that the response is zero for $t<0$. The special form of the impulse response for the $(0,0,0)$ mode is a consequence of the associated natural frequency being zero. The corresponding representation of the pressure at an arbitrary field point is found from Eq. (10.3.10) to be

$$
\begin{equation*}
p(\bar{x}, t)=\frac{4 \rho_{0} c^{2} q_{0}}{L H} \sum_{j=0}^{J} \sum_{n=0}^{N}\left(B_{j, n, 0}\right)^{2} \cos \left(\frac{j \pi x}{L}\right) \cos \left(\frac{n \pi y}{H}\right) g_{j, n, 0}(t) \tag{7}
\end{equation*}
$$

The dimensionality of $q_{0}$ is (volume/time)/length $=$ length ${ }^{2} /$ time. Therefore, dividing $p$ by $\rho_{0} c q_{0} / L$ is a convenient way of eliminating dependence on the unspecified value of $q_{0}$.

An important aspect is truncation of the series. The impulse response function $g_{j, n, 0}(t)$ is the only factor in Eq. (7) whose magnitude varies significantly from term to term. Its magnitude is proportional to $1 / \omega_{j, n}$, so it decreases monotonically as $j$ or $n$ increases. Thus, truncation will be based in assuring that any frequencies that are omitted exceed some maximum value. Let us describe this frequency limit as $\sigma \min \left(\omega_{\text {nat }}\right)$, where $\sigma$ is a factor to be determined. Because $L>H$, the fundamental nonzero frequency is $\omega_{1,0,0}=\pi c / L$. (The zero frequency is not relevant.) Thus, the task is to identify values of $J$ and $N$ such that $\omega_{J, n, 0}$, for any $n$, and $\omega_{j, N, 0}$, for any $j$, exceed $\sigma(\pi c / L)$. The largest $J$ corresponds to $n=0$, so we want $\omega_{J, 0,0} \geq \sigma(\pi c / L)$, that is $J \geq \operatorname{ceil}(\sigma)$. A similar analysis sets $N \geq \operatorname{ceil}(\sigma H / L)$.

A matrix algorithm may be implemented to evaluate Eq.(7) at specified locations over any time interval. The values of $J$ and $N$ will ultimately be found to be quite large, which means that there are many modes to compute. There are several ways in which the algorithm in Example 10.2 may be adapted to the present formulation. The following was found to be quite efficient relative to others that were tried, yet able to run on a laptop whose memory workspace is limited. Variants that work on more powerful computers might be more rapid, but their workspace requirements could be excessive. The first step creates a row vector of mode numbers $[m]=\left[\begin{array}{llll}1 & 2 & 3 & \ldots M\end{array}\right]$, where $M=(J+1)(N+1)$. The pair of wavenumbers in the $x$ and $y$ direction are placed in row vectors [ $j$ ] and [n], whose length is the same length as $[m]$. Operations to create these quantities are $[j]=$ floor $(([m]-1) /(N+1))$ followed by $[n]=[m]-1-[j] *(N+1)$. These row vectors are used to vectorize computation of the natural frequencies in Eq. (2) and normalizing coefficients in Eq. (3). The MATLAB code for these operations is omega=c*sqrt((j*pi/L).^2+(n*pi/H).^2); B=2./sqrt((1+
$(j==0)) . *(1+(n==0)))$; The result is row vectors $[\omega]$ and $[B]$ whose size matches $[m]$. To evaluate the modal values, let $\bar{x}_{f}, f=1, \ldots F$, denote the field points at which we wish to evaluate the pressure, and let $t_{\ell}$ be a sequence of instants at which the pressure is to computed. The values of all mode functions at fixed $x$ are stored in row $f$ of a rectangular array $[\chi]$. Pressure values at several locations may be computed by incrementing $f$, thereby stacking rows of $[\chi]$. In MATLAB, the vectorized implementation of this operation defines two vectors $\mathrm{x} \_\mathrm{f}$ and $\mathrm{Y} \_\mathrm{f}$ that hold the coordinates of the set of $\bar{x}_{f}$ locations. Then, Chi (f,:) =B.^ 2.* $\cos \left(j * p i * x \_f / L\right) .{ }^{*} \cos \left(n * p i * Y \_f / H\right) ;$.
The pressure in Eq. (7) at the various locations at a specific instant $t_{\ell}$ may be computed as the product of $[\chi]$ and a column vector of $g_{m} \equiv g_{j, m, 0}$ values, that is,

$$
\begin{equation*}
\left\{p\left(\{\bar{x}\}, t_{s}\right)\right\}=\frac{\rho_{0} c q_{0}}{L} \frac{4 c}{H}[\chi]\left\{g\left(t_{s}\right)\right\} \tag{8}
\end{equation*}
$$

A time history may be constructed by adjoining column vectors of $\{p\}$ values at each instant, that is,

$$
\left[\left\{p\left(\{\bar{x}\}, t_{1}\right)\right\}\left\{p\left(\{\bar{x}\}, t_{2}\right)\right\} \cdots\right]=\frac{\rho_{0} c q_{0}}{L} \frac{4 c}{H}[\chi]\left[\left\{g\left(t_{1}\right)\right\}\left\{g\left(t_{2}\right)\right\} \cdots\right]
$$

The highest frequency sets the time step according to the Nyquist criterion. It states that the sampling interval should be less than half the period of the highest frequency contained in a signal. The modal series has been truncated $J$ and $N$. This sets the highest natural frequency at $\omega_{J, N, 0}$. The Nyquist criterion is $\Delta t \leq 0.5\left(2 \pi / \omega_{J, N, 0}\right)$. The system parameters, $c=340 \mathrm{~m} / \mathrm{s}, L=4 \mathrm{~m}, H=3 \mathrm{~m}$, give $\omega_{J, M, 0}=377\left(10^{3}\right)$ $\mathrm{rad} / \mathrm{s}$, which leads to $\Delta t \leq 8.32 \mu \mathrm{~s}$.

For the sake of completeness, the signals received at three points along a diagonal were computed: $(x, y)=(L / 4,3 H / 4),(L / 2, H / 2)$, and $(3 L / 4, H / 4)$. The factor for modal truncation was set at $\sigma=100$, which leads to $J=100, N=75$. The result appears in Fig. 2.


Figure 2.
The plots have a jagged appearance. Calculations with $\Delta t$ set at one-tenth of the Nyquist value primarily smoothed out the fluctuations, whereas increasing the value of $\sigma$ led to curves whose shape changed from Fig. 2. In other words, a modal truncation based on $\sigma=100$ is not adequate. All computations were repeated with $\sigma=1000$. This leads to $J=1000, N=750$, which corresponds to 751751 modes.

The sampling interval under the Nyquist sampling guideline is $\Delta t=0.832 \mu \mathrm{~s}$. Raising the number of modes and number of instants by a factor of ten relative to Fig. 2 leads to a factor of 100 increase in the computational time. Nevertheless, the effort is still manageable, with an elapsed time of 1.5 min on a laptop computer with an Intel I7 CPU.

The result of the refined computation is the data plotted in Fig.3. Each plot is much smoother. Some high-frequency wiggles are present. For reasons elucidated below, it is not possible to eliminate them. At each location described in Fig. 3, the pressure shows a sharp peak, followed by a rapid decay. The shape of each signal after this arrival time is reminiscent of the solution to Example 6.5, which explored the field radiated by an impulsive line source in free space. Indeed, that solution can be used in conjunction with the method of images to construct the response of the present system.


Figure 3.
According to the method of images the walls at constants $x$ and $y$ serve as mirrors for the line source at $x=y=0$. Figure 4 shows that an infinite number of image sources is required to satisfy the condition that $\bar{n} \cdot \nabla p=0$ at all walls. A subtle aspect of the image construction is that the walls at $z=0$ and $z=W$ also serve as mirrors. This leads to a continuous line extending to infinity in the positive and negative $z$ direction. In other words, each image is an impulsive infinite line source, each of whose mass acceleration per unit length is $4 \rho_{0} q_{0} \delta(t)$. Confirmation of the correctness of this model is the fact that the time at which each peak occurs in Fig. 3 is the distance from a field point to one of the images in Fig. 4, divided by $c$. This is the time required for a signal leaving the image line source at $t=0$ to arrive at the field point.


Figure 4.
A corollary is that the waveform received at any point within the cavity begins with arrival of the signal that travels directly from the source, followed by the
arrival of pulses resulting from one or more reflections of the source signal. As the distance along the reflected path increases, the arriving signal is weakened by cylindrical spreading, as explained in Example 6.5. This conceptual picture explains the waveforms in Fig. 5, which were obtained by halting computation of Eq. (8) at $t_{\max }=3 L / c$. Multiple pulses are seen at each field point. The arrival time for some is sufficiently close that they overlap.


The availability of an alternate solution makes it possible to recognize some faults with the cavity mode series solution. An exceptionally large number of modes must be included in the series in order to obtain a high-quality waveform. In contrast, the method of images merely requires inclusion of only those images whose signal arrives at the selected field point within the designated time interval. The field of an impulsive line source is singular at the first arrival time. The modal solution cannot replicate this singularity. Furthermore, the modal solution shows high-frequency wiggles prior to the arrival of the signal from each image. This is an impossibility, because early arrival corresponds to a signal that travels supersonically from an image to the field point. Both artifices stem from the fact that the impulsive line source excites all modes equally, as is evident in Eq. (4) for the $S_{j, n}$ values. A very large number of modes is required to reproduce the discontinuity. Increasing the number of modes as we did decreases the scale of the precursor signal in duration and amplitude. However, the precursor cannot be eliminated, and the singularity cannot be replicated, with a finite length modal series.

### 10.3.5 Cylindrical Cavity

The analysis of cylindrical cavity modes proceeds much like the development for rectangular cavities. We align the $x$-axis with the centerline, with the origin placed at an end. The mode functions we seek are solutions of the homogeneous Helmholtz equation subject to rigid or pressure-release conditions at the ends $x=0$ and $x=L$, and on the interior wall of the cylinder, $R=a$. In addition, the mode function must be finite at $R=0$. This excludes the Neumann function as a possible description of the transverse dependence, so that dependence will be a Bessel function. (If there is a concentric inner cylinder, the finiteness condition is replaced by a boundary
condition on the inner cylinder. In that case, a mode function is formed from both the Bessel and Neumann functions.) We designate the transverse wavenumber as $\mu$. The behavior in the axial direction $x$ is like that between parallel walls in a rectangular waveguide. If the field is not axisymmetric, then the dependence on the azimuthal angle $\theta$ must be harmonic. Hence, any cavity mode may be described as

$$
\begin{equation*}
\Psi=B J_{n}(\mu R) F(n \theta)\left[c_{x} \cos (\kappa x)+s_{x} \sin (\kappa x)\right] \tag{10.3.24}
\end{equation*}
$$

where

$$
\begin{equation*}
F(n \theta)=\cos (n \theta), \sin (n \theta), \text { or } \exp ( \pm i n \theta) \tag{10.3.25}
\end{equation*}
$$

Setting $P=0$ or $\partial P / \partial x=0$ at $x=0$ leads to either $c_{x}=0$ or $s_{x}=0$. Then, satisfying $P=0$ or $\partial P / \partial x=0$ at $x=L$ yields an infinite number of $\kappa L$ values. The possible values of $\mu a$ are determined by satisfying the rigid or pressure-release conditions at $R=a$. This leads to the $\mu a$ values being either a zero of a Bessel function, or a value at which a Bessel function has an extreme value. The specific conditions are

$$
\text { Rigid: }\left.\frac{\partial \Psi}{\partial R}\right|_{R=a}=0 \Longrightarrow J_{n}^{\prime}\left(\eta_{n, m}\right)=0, \mu=\frac{\eta_{n, m}}{a}, m=1,2, \ldots
$$

$$
\begin{equation*}
\text { Pressure-release: }\left.\Psi\right|_{R=a}=0 \Longrightarrow J_{n}\left(\eta_{n, m}\right)=0, \mu=\frac{\eta_{n, m}}{a}, \quad m=1,2, \ldots \tag{10.3.26}
\end{equation*}
$$

Thus, three subscripts are required to denote a cylindrical cavity mode. The first is the order $n$ of the Bessel function, which is the same as the azimuthal harmonic $n$. The second is the root number $m$ for the radial wavenumber. The third is the root number $j$ for the axial wavenumber $\kappa$. Furthermore, as indicated by Eq. (10.3.25), two independent functions describe the dependence on $\theta$ for a given triad of $n, m, j$ values. Cavity modes are standing waves, so the sinusoidal functions are more appropriate in the present context. We shall use a subscript $\alpha=c$ or $\alpha=s$ to indicate which sinusoidal function is being considered, with this subscript omitted for the axisymmetric mode, $n=0$. Correspondingly, the modes $\Psi_{n, m, j, \alpha}$ of a cylindrical cavity are

$$
\begin{gather*}
\Psi_{0, m, j}=B_{n, m, 0} J_{0}\left(\eta_{0, m} \frac{R}{a}\right)\left[c_{x} \cos \left(\kappa_{j} x\right)+s_{x} \sin \left(\kappa_{j} x\right)\right] \\
\Psi_{n, m, j, c}=B_{m, n, j} J_{n}\left(\eta_{n, m} \frac{R}{a}\right)\left[c_{x} \cos \left(\kappa_{j} x\right)+s_{x} \sin \left(\kappa_{j} x\right)\right] \cos (n \theta), n>0 \\
\Psi_{n, m, j, s}=B_{m, n, j} J_{n}\left(\eta_{n, m} \frac{R}{a}\right)\left[c_{x} \cos \left(\kappa_{j} x\right)+s_{x} \sin \left(\kappa_{j} x\right)\right] \sin (n \theta), n>0 \tag{10.3.27}
\end{gather*}
$$

It still remains to determine the natural frequency. This is done by ascertaining the frequency at which a mode function satisfies the Helmholtz equation,

$$
\begin{equation*}
\frac{\partial^{2} P}{\partial R^{2}}+\frac{1}{R} \frac{\partial P}{\partial R}+\frac{1}{R^{2}} \frac{\partial^{2} P}{\partial \theta^{2}}+\frac{\partial^{2} P}{\partial z^{2}}+k^{2} P=0 \tag{10.3.28}
\end{equation*}
$$

We also have $\partial^{2} \Psi_{n, m, j, \alpha} / \partial \theta^{2}=-n^{2} \Psi_{n, m, j, \alpha}$ and $\partial^{2} \Psi_{n, m, j, \alpha} / \partial x^{2}=-\kappa_{j}^{2} \Psi_{n, m, j, \alpha}$. In the transverse direction, we know that $J_{n}\left(\eta_{n, m} R / a\right)$ is a solution of Bessel's equation, which means that

$$
\begin{equation*}
\frac{\partial^{2} \Psi_{n, m, j, \alpha}}{\partial R^{2}}+\frac{1}{R} \frac{\partial \Psi_{n, m, j, \alpha}}{\partial R}=-\left(\frac{\eta_{n, m}^{2}}{a^{2}}-\frac{n^{2}}{R^{2}}\right) \Psi_{n, m, j, \alpha} \tag{10.3.29}
\end{equation*}
$$

It follows that $\Psi_{n, m, j}$ is a solution of the Helmholtz equation if

$$
\begin{equation*}
\omega_{n, m, j, \alpha} \equiv c k_{j, m, n}=c\left(\kappa_{j}^{2}+\frac{\eta_{n, m}^{2}}{a^{2}}\right)^{1 / 2} \tag{10.3.30}
\end{equation*}
$$

The azimuthal harmonic number $n$ does not appear explicitly in this expression. Nevertheless, inclusion of $n$ in the list of subscripts for the natural frequency is warranted by the fact that the transverse wavenumbers $\eta_{n, m}$ depend on which Bessel function order is under consideration.

An important aspect of the nonaxisymmetric $(n>0)$ modes is that they occur as a pair $\alpha=c$ and $\alpha=s$ with equal natural frequencies. Despite the equality of natural frequencies, each function is orthogonal to all others. The specific forms of the general modal orthogonality properties and normalization definition, Eqs. (10.3.8) and (10.3.9), for a cylindrical cavity are

$$
\begin{gather*}
\int_{0}^{L} \int_{-\pi}^{\pi} \int_{0}^{a} \Psi_{n, m, j, \alpha} \Psi_{r, s, \ell, \beta} R d R d \theta d x=\delta_{n, r} \delta_{m, s} \delta_{j, \ell} \delta_{\alpha, \beta}\left(\pi a^{2} L\right) \\
\int_{0}^{L} \int_{-\pi}^{\pi} \int_{0}^{a}\left(\frac{\partial \Psi_{n, m, j, \alpha}}{\partial x} \frac{\partial \Psi_{r, s, i, \beta}}{\partial x}+\frac{\partial \Psi_{n, m, j, \alpha}}{\partial R} \frac{\partial \Psi_{r, s, i, \beta}}{\partial R}\right. \\
\left.+\frac{1}{R^{2}} \frac{\partial \Psi_{n, m, j, \alpha}}{\partial \theta} \frac{\partial \Psi_{r, s, i, \beta}}{\partial \theta}\right) R d R d \theta d z=\delta_{n, r} \delta_{m, s} \delta_{j, \ell} \delta_{\alpha, \beta}\left(\omega_{j, m, n, \alpha}\right)^{2}\left(\pi a^{2} L\right) \tag{10.3.31}
\end{gather*}
$$

There is no need to prove these properties because they are covered by the general proof.

The case where all mode numbers match, that is, $n=r, m=s, j=\ell$, and $\alpha=\beta$, is used to set the normalization coefficients $B_{n, m, j}$. Note that the same $B_{n, m, j}$ value applies to the cosine and sine modes because one is merely rotated about the $x$-axis relative to the other. An actual determination of the $B_{n, m, j}$ values will entail evaluation of the integral of a product of Bessel functions. Because $\eta_{n, m}$ as an eigenvalue, a formula ${ }^{7}$ is available for this integral. The result depends on whether the surface is pressure-release or rigid, specifically

[^39]$\int_{0}^{a} J_{n}\left(\eta_{n, m} \frac{R}{a}\right)^{2} R d R=\left\{\begin{array}{l}\frac{a^{2}}{2} J_{n}^{\prime}\left(\eta_{n, m}\right)^{2} \text { if } J_{n}\left(\eta_{n, m}\right)=0 \\ \frac{a^{2}}{2}\left[1-\left(\frac{n}{\eta_{n, m}}\right)^{2}\right] J_{n}\left(\eta_{n, m}\right)^{2} \text { if } J_{n}^{\prime}\left(\eta_{n, m}\right)=0 \text { and } \eta_{n, m} \neq 0\end{array}\right.$
Upon determination of the mode functions and natural frequencies, an analysis of response in the time domain or frequency domain may proceed by following the steps in the general development provided in Sect. 10.3.1.

EXAMPLE 10.9 A box and a cylinder contain air at standard conditions. The cross section of the box is square, with a side length of $w$, and the radius of the cylinder is $a$. Both have equal length $L$ in the axial direction, which is designated as $x$. The ends at $x=0$ and $x=L$ are open to the atmosphere and each cross section is sufficiently large to set $P=0$ at both ends. The other sides are rigid. The parameter that is of interest is the smallest natural frequency corresponding to a mode that is not planar transverse to the $x$ axis. In particular, it is desired that this frequency be a specified value $\Omega$ for both configurations. Derive expressions for $a$ and $w$ for which this condition occurs.

## Significance

This example will provide quantitative insight into the manner in which dimensions and shape affect natural frequencies.

## Solution

Let us begin with a circular cross section. The ends are pressure-release, and the same frequency results from using $\cos (n \theta)$ and $\sin (n \theta)$, so a mode function is described by

$$
\Psi_{n, m, j, c}=B J_{n}\left(\mu_{n, m} R\right) \cos (n \theta) \sin \left(\frac{j \pi x}{L}\right)
$$

where $j$ is the root number for the axial eigenvalue, $n$ is the azimuthal harmonic, and $m$ is the root number for the transverse eigenvalue. The latter comes from satisfying the condition $V_{R}=0$ at $R=a$, for which the characteristic equation is

$$
J_{n}^{\prime}\left(\eta_{n, m}\right)=0, \mu_{n, m}=\frac{\eta_{n, m}}{a}
$$

Equation (10.3.30) tells us that the lowest frequency corresponds to minimum values of $\mu_{n, m}$ and $\kappa_{j}$. To identify the appropriate value of the former, we examine Fig. 1, which is a graph of the Bessel function at various orders. The roots $\eta_{n, m}$ are the argument $x$ for which $J_{n}(x)$ has a maximum or minimum. The first maximum occurs for $n=0$, but this root gives $J_{0}(0)=1$, which is a planar mode. The smallest nonzero value of $x$ leading to an extreme value of $J_{n}(x)$ occurs at $x \approx 1.8$ for $n=1$. A more accurate value may be found by invoking the identity $J_{1}^{\prime}(\xi)=J_{0}(x)-J_{1}(x) / x$ to implement a numerical solution. Doing so yields $\eta_{1,1}=1.8412$.


Figure 1.

The axial wavenumber is $\kappa_{j}=j \pi / L$, with $j>0$. The minimum occurs for $j=1$. Thus, the fundamental frequency for nonplanar modes is found from Eq. (10.3.30) to be

$$
\min \left(\omega_{j, m, n}\right)=\omega_{1,1,1}=c\left[\left(\frac{\pi}{L}\right)^{2}+\left(\frac{\eta_{1,1}}{a}\right)^{2}\right]^{1 / 2}
$$

The $x$ dependence of the modes of the box is the same as it is for the cylinder. The dependence on $y$ and $z$ must be harmonic, with zero slope at the walls. The corresponding modes are

$$
\Psi_{j, m, n}=B \cos \left(\frac{m \pi y}{w}\right)\left(\frac{n \pi z}{w}\right) \sin \left(\frac{j \pi x}{L}\right), m=0,1, \ldots, \quad n=0,1, \ldots
$$

The natural frequencies are

$$
\omega_{j, m, n}=c\left[\left(\frac{j \pi}{L}\right)^{2}+\left(\frac{m \pi}{w}\right)^{2}+\left(\frac{n \pi z}{w}\right)^{2}\right]^{1 / 2}
$$

The lowest natural frequency corresponds to the minimum values of $j, n$, and $m$. The minimum value of $j$ is one. Although the smallest values of $m$ and $n$ are zero, both being zero lead to a planar mode. However, setting either $m$ or $n$ to zero and the other to one is acceptable. Thus, we find that

$$
\min \left(\omega_{j, m, n}\right)=\omega_{1,1,0}=\omega_{1,0,1}=c\left[\left(\frac{\pi}{L}\right)^{2}+\left(\frac{\pi}{w}\right)^{2}\right]^{1 / 2}
$$

The minimum for each configuration should be $\Omega$, so

$$
\left(\frac{\Omega}{c}\right)^{2}=\left(\frac{\pi}{L}\right)^{2}+\left(\frac{\eta_{1,1}}{a}\right)^{2}=\left(\frac{\pi}{L}\right)^{2}+\frac{\pi^{2}}{w^{2}}
$$

This yields

$$
a=\frac{1.8412}{\left[\left(\frac{\Omega}{c}\right)^{2}-\left(\frac{\pi}{L}\right)^{2}\right]^{1 / 2}}, \quad w=\frac{\pi}{\left[\left(\frac{\Omega}{c}\right)^{2}-\left(\frac{\pi}{L}\right)^{2}\right]^{1 / 2}}
$$

Although both dimensions depend on $L$ and $\Omega / c$, their ratio is a constant,

$$
\frac{a}{w}=\frac{1.8412}{\pi}
$$

The corresponding ratio of the cross-sectional areas is

$$
\frac{\mathcal{A}_{\mathrm{cyl}}}{\mathcal{A}_{\mathrm{rect}}}=\frac{\pi a^{2}}{w^{2}}=1.0791
$$

The areas are quite close. The associated mode functions also have some attributes in common. Figure 2 shows equal value contours of the modes at $x=L / 2$. The thickened contours are nodal lines. The $(1,1,1)$ mode with a $\cos \theta$ dependence and the $(1,1,0)$ mode of the box are symmetric with respect to the horizontal centerline and antisymmetric with respect to the vertical centerline. Their only nodal line is along the vertical centerline. The patterns are rotated $90^{\circ}$ for the $(1,1,1, \cos \theta)$ mode of a cylinder and the $(1,0,1)$ mode of a box.


Figure 2.

### 10.3.6 Spherical Cavity

The objective here is to identify the axisymmetric natural modes of a spherical cavity whose wall is either rigid or pressure-release. In general, the individual terms in a series of forced modes satisfy the Helmholtz equation. Equation (10.2.14) is such a series for an axisymmetric field within a sphere. The individual terms in that series are spherical harmonics. The terms are summed in order to satisfy the boundary condition of an arbitrary excitation on the spherical wall. An analysis of natural cavity modes instead requires that each spherical harmonic individually be consistent with either a zero radial velocity or zero pressure at the boundary. Thus, we begin by considering a natural mode whose form is

$$
\begin{equation*}
\Psi=B j_{m}(k r) P_{m}(\cos \psi) \tag{10.3.33}
\end{equation*}
$$

Although the spherical Neumann function also satisfies the Helmholtz equation, it is excluded by the condition that the pressure at the center must be finite.

As in the previous analyses, the $B$ coefficient will be set by the orthogonality property. The only other parameter not yet set is $k$, which is the ratio of the natural frequency to the speed of sound. This quantity is found by satisfying the condition at the wall. Rigidity there requires that $\partial \Psi / \partial r=0$ at $r=a$, whereas a pressure-release condition requires that $\Psi=0$ at $r=a$. The characteristic equations that result from satisfying these alternative conditions are

$$
\left.\begin{array}{c}
\text { Rigid: } j_{m}^{\prime}\left(\eta_{m, n}\right)=0  \tag{10.3.34}\\
\text { Pressure-release: } j_{m}\left(\eta_{m, n}\right)=0
\end{array}\right\}, \quad \eta_{m, n} \equiv k_{m, n} a \equiv \omega_{n, m} \frac{a}{c}
$$

The $\eta$ quantities, which are the eigenvalues, have two subscripts. The first, $n$, is the polar harmonic number. Because the spherical Bessel function is oscillatory about a zero value, there are multiple roots at a specified $n$. The second subscript, $m$, is the root number. Usually, the roots and therefore the natural frequencies are sequenced in ascending order, so that $\eta_{m, n}<\eta_{m, n+1}$.

Each eigenvalue might lead to a different normalizing coefficient, so those values are denoted as $B_{n, m}$. Thus, we have established that the axisymmetric natural modes of any spherical cavity are

$$
\begin{equation*}
\Psi_{n, m}=B_{n, m} j_{m}\left(\eta_{n, m} \frac{r}{a}\right) P_{n}(\cos \psi), \quad \omega_{n, m}=\frac{c}{a} \eta_{n, m} \tag{10.3.35}
\end{equation*}
$$

Despite the common mathematical form of the mode functions, their radial derivative will vanish at $r=a$ for a rigid wall, whereas they will be zero at that location for a pressure-release wall. Because $j_{0}(0)=1$, the first root for $m=0$ in the rigid wall case is $\eta_{0,0}=0$. Such a mode corresponds to a uniform pressure distribution in the transverse direction. All other eigenvalues are positive.

The first few roots for each case are given in Table 10.1. As is usual for Bessel functions, the higher roots are well approximated by adding multiples of $\pi$ to these values.

Table 10.1 Nondimensional natural frequencies for the axisymmetric modes of a spherical cavity whose wall is either rigid or pressure-release

| Wall condition | Polar harmonic (m) | $k_{m, 1} a$ | $k_{m, 2} a$ | $k_{m, 3} a$ | $k_{m, 4} a$ |
| :--- | :--- | :--- | :--- | :--- | :--- |
| Rigid | 0 | 0 | 4.4934 | 7.7253 | 10.9041 |
|  | 1 | 2.0816 | 5.9404 | 9.2058 | 12.4044 |
|  | 2 | 3.3421 | 7.2899 | 10.6139 | 13.8461 |
| Pressure-release | 0 | 3.1416 | 6.2832 | 9.4248 | 12.5664 |
|  | 1 | 4.4934 | 7.7253 | 10.9041 | 14.0662 |
|  | 2 | 5.7635 | 9.0950 | 12.3229 | 15.5146 |

Example 10.7 showed that the modal density of a rectangular cavity increases as the range of natural frequencies is increased. The spherical Bessel functions are not exactly periodic functions, so there will not be cases where the eigenvalues for different modes are identical. However, they are nearly periodic for large arguments, so modes having different harmonic numbers $n$ might have very close natural frequencies.

The orthogonality conditions for the spherical cavity modes are obtained by specializing the general conditions, Eqs. (10.3.8) and (10.3.9), to an axisymmetric spherical geometry. Thus, we use a differential volume element in the shape of a ring whose radius is $r \sin \psi$, that is, $d V=(2 \pi r \sin \psi)(r d \psi) d r$. The result is

$$
\begin{gather*}
2 \pi \int_{0}^{a} \int_{0}^{\pi} \Psi_{m, n} \Psi_{r, s} r^{2} \sin \psi d \psi d r=\frac{4}{3} \pi a^{3} \delta_{m, r} \delta_{n, s} \\
2 \pi \int_{0}^{a} \int_{0}^{\pi}\left[\frac{\partial \Psi_{m, n}}{\partial r} \frac{\partial \Psi_{r, s}}{\partial r}+\frac{1}{r^{2}} \frac{\partial \Psi_{m, n}}{\partial \psi} \frac{\partial \Psi_{r, s}}{\partial \psi}\right] r^{2} \sin \psi d \psi d r=\frac{4 \pi}{3} a^{3} \delta_{m, r} \delta_{n, s} \omega_{m, n}^{2} \tag{10.3.36}
\end{gather*}
$$

As was done for other geometries, the case where the mode numbers match, $m=r$ and $n=s$ are used to set the scaling coefficients $B_{m, n}$. The first of the above equations is used for this purpose. The actual evaluation will entail evaluation of an integral containing products of Bessel functions. We may obtain a formula by introducing into a tabulated integral, ${ }^{8}$ the relation of spherical and cylindrical Bessel functions, specifically,

$$
\begin{equation*}
j_{n}\left(\eta_{n, m} \frac{r}{a}\right)=\left(\frac{\pi a}{2 \eta_{n, m}}\right)^{1 / 2} J_{n+1 / 2}\left(\eta_{n, m} \frac{r}{a}\right) \tag{10.3.37}
\end{equation*}
$$

This relation is useful for computations, as well as for the normalization integrals. For the latter, it leads to

[^40]\[

\int_{0}^{a} j_{n}\left(\eta_{n, m} \frac{r}{a}\right)^{2} r^{2} d r=\left\{$$
\begin{array}{l}
\left\{\frac{1}{2} a^{3} j_{n+1}\left(\eta_{n, m}\right)^{2} \text { if } j_{n}\left(\eta_{n, m}\right)=0\right.  \tag{10.3.38}\\
\frac{1}{2} a^{3}\left[1-\left(\frac{n+1 / 2}{\eta_{n, m}}\right)^{2}\right] j_{n}\left(\eta_{n, m}\right)^{2} \text { if } j_{n}^{\prime}\left(\eta_{n, m}\right)=0 \\
\text { and } \eta_{n, m}>0
\end{array}
$$\right\}
\]

The features unique to the spherical geometry are the specific form of the mode functions, the natural frequencies, and evaluation of the normalizing coefficients. After they have been determined, analysis of the pressure field may follow the general path described in Sect. 10.3.1.


#### Abstract

EXAMPLE 10.10 A spherical tank used for chemical processing consists of concentric spheres that separate different fluids. The outer sphere, whose radius is $a$, is rigid. The inner sphere, whose radius is $b$, is a permeable membrane that allows the fluids to mix slowly by diffusing through it. This sphere is well approximated as having zero thickness with neither inertia nor stiffness. The density and sound speed within the inner sphere are $\rho_{1}$ and $c_{1}$, while $\rho_{2}$ and $c_{2}$ are the properties of the fluid contained between the inner and outer spheres. Derive the characteristic equation and an expression for the cavity modes of this system. Also identify the orthogonality conditions for these modes.


## Significance

Our study of propagation through planar layers entailed making solutions of the Helmholtz equation satisfies continuity conditions at interfaces. Those concepts are extended here to spherical waves.

## Solution

It is stated that the permeable membrane physically separates the inner and outer regions, but cannot resist any movement. Thus, it must be that the pressure and particle velocity on either side of the inner sphere are equal. In addition, the pressure in both regions must satisfy the Helmholtz equation. The definition of a mode requires that all locations in a system exhibit an oscillation at the same frequency. Thus, the acoustic wavenumbers for the Helmholtz equation are $k_{1}=\omega c_{1}$ and $k_{2}=\omega c_{2}$. The other boundary conditions to be met are zero radial velocity at $r=a$ and finiteness of $P$ at $r=0$. Hence, we seek pressure fields $P^{(1)}$ and $P^{(2)}$ that satisfy

$$
\begin{gathered}
\nabla^{2} P^{(1)}+\left(k_{1}\right)^{2} P^{(1)}=0, \quad 0<r<b \\
\nabla^{2} P^{(2)}+\left(k_{2}\right)^{2} P^{(2)}=0, \quad b<r<a \\
\frac{\partial P^{(2)}}{\partial r}=0 \text { at } r=b, \quad P^{(1)} \text { is finite at } r=0 \\
\frac{1}{i \omega \rho_{1}} \frac{\partial P^{(1)}}{\partial r}=P^{(2)} \text { at } r=a \\
i \omega \rho_{2} \\
\frac{\partial P^{(2)}}{\partial r} \text { at } r=b
\end{gathered}
$$

The continuity conditions at $r=b$ require that the same polar harmonic $P_{n}(\cos \Psi)$ be associated with the mode in the inner and outer regions. Finiteness at $r=0$ requires that the radial dependence of $P^{(1)}$ be a spherical Bessel function. The outer region does not contain $r=0$, so both the spherical Bessel and Neumann functions are required to describe the radial dependence there. It follows that a suitable ansatz for the pressure in each region is

$$
\begin{gather*}
P^{(1)}=B^{(1)} j_{n}\left(k_{1} r\right) P_{n}(\cos \Psi) \\
P^{(2)}=\left[B^{(2)} j_{n}\left(k_{2} r\right)+C^{(2)} n_{n}\left(k_{2} r\right)\right] P_{n}(\cos \Psi) \tag{1}
\end{gather*}
$$

When these expressions are substituted into the boundary and continuity conditions, the Legendre function is a common factor for each, so it may be canceled. The resulting equations are

$$
\begin{gathered}
k_{2}\left[B^{(2)} j_{n}^{\prime}\left(k_{2} a\right)+C^{(2)} n_{n}^{\prime}\left(k_{2} a\right)\right]=0 \\
B^{(1)} j_{n}\left(k^{(1)} b\right)=B^{(2)} j_{n}\left(k_{2} b\right)+C^{(2)} n_{n}\left(k_{2} b\right) \\
\left(\frac{1}{\rho_{1} c_{1}}\right) B^{(1)} j_{n}^{\prime}\left(k_{1} b\right)=\left(\frac{1}{\rho_{2} c_{2}}\right)\left[B^{(2)} j_{n}^{\prime}\left(k_{2} b\right)+C^{(2)} n_{n}^{\prime}\left(k_{2} b\right)\right]
\end{gathered}
$$

The first equation is satisfied by $k_{2}=0$, which leads to $\omega=0$. Because the outer boundary is rigid, this is the zero frequency mode. The general discussion suggests that such a mode exists for this system. This will be proven after we find the modes whose natural frequency is greater than zero. Thus, $k_{2}$ may be factored out of the first equation.

It is convenient to express the coefficient equations in matrix form. In recognition that the frequency sets the acoustic wavenumbers, we replace $k_{1}$ and $k_{2}$ with $\omega / c_{1}$ and $k_{2}=\omega / c_{2}$, respectively. Hence, the matrix representation of the coefficient equations is

$$
\begin{align*}
& {[D(\omega)]\left[\begin{array}{lll}
B^{(1)} & B^{(2)} & C^{(2)}
\end{array}\right]^{\mathrm{T}}=\{0\}} \\
& {[D(\omega)]=\left[\begin{array}{ccc}
0 & j_{n}^{\prime}\left(\omega a / c_{2}\right) & n_{n}^{\prime}\left(\omega a / c_{2}\right) \\
j_{n}\left(\omega b / c_{1}\right) & -j_{n}\left(\omega b / c_{2}\right) & -n_{n}\left(\omega b / c_{2}\right) \\
\left(\frac{\rho_{2} c_{2}}{\rho_{1} c_{1}}\right) j_{n}^{\prime}\left(\omega b / c_{1}\right) & -j_{n}^{\prime}\left(\omega b / c_{2}\right) & -n_{n}^{\prime}\left(\omega b / c_{2}\right)
\end{array}\right]} \tag{2}
\end{align*}
$$

At an arbitrary $\omega$, the solution of these equations is $B^{(1)}=B^{(2)}=C^{(2)}=0$. A nontrivial solution requires that $\omega$ be a value that causes the determinant of $[D(\omega)]$ to be zero, thereby making it impossible to invert the matrix. Thus, the characteristic equation is

$$
\begin{equation*}
|[D(\omega)]|=0 \tag{3}
\end{equation*}
$$

The roots of this (scalar) equation constitute the eigenvalues. Stated differently, they are the natural frequencies corresponding to the axisymmetric cavity modes. We could reduce the determinant to an algebraic expression, but there is little reason to do so because evaluation of its roots would require numerical methods. As was
the case for a homogeneous cavity, the oscillatory nature of the Bessel and Neumann functions leads to an infinite number of natural frequencies $\omega_{n, m}$ for a specified polar harmonic $n$.

When the determinant of [ $D$ ] vanishes, one of the coefficient equations described in Eq. (2) is not independent of the other two. If we drop the third equation, we may solve the first two equations for $B^{(2)}$ and $C^{(2)}$ as ratios to $B^{(1)}$, that is

$$
\begin{gathered}
B^{(2)}=\beta_{n, m} B^{(1)}, \quad C^{(2)}=\gamma_{n, m} B^{(1)} \\
\beta_{n, m}=\frac{j_{n}\left(\omega_{n, m} b / c_{1}\right) n_{n}^{\prime}\left(\omega_{n, m} a / c_{2}\right)}{j_{n}\left(\omega_{n, m} b / c_{2}\right) n_{n}^{\prime}\left(\omega_{n, m} q / c_{2}\right)-n_{n}\left(\omega_{n, m} b / c_{2}\right) j_{n}^{\prime}\left(\omega_{n, m} a / c_{1}\right)} B^{(1)} \\
\gamma_{n, m}=-\frac{j_{n}\left(\omega_{n, m} b / c_{1}\right) j_{n}^{\prime}\left(\omega_{n, m} a / c_{2}\right)}{j_{n}\left(\omega_{n, m} b / c_{2}\right) n_{n}^{\prime}\left(\omega_{n, m} q / c_{2}\right)-n_{n}\left(\omega_{n, m} b / c_{2}\right) j_{n}^{\prime}\left(\omega_{n, m} a / c_{1}\right)} B^{(1)}
\end{gathered}
$$

The coefficient $B^{(1)}$ is set by the normalization part of the orthogonality condition, so it also will depend on $n$ and $m$. Therefore, we have established that a mode is described by

$$
\begin{equation*}
\Psi_{n, m}=B_{n, m}^{(1)} F_{n, m}(r) P_{n}(\cos \Psi) \tag{4}
\end{equation*}
$$

where the radial function is

$$
F_{n, m}(r)=\left\{\begin{array}{l}
j_{n}\left(\omega_{n, m} r / c_{1}\right), \quad 0<r<b  \tag{5}\\
\beta_{n, m} j_{n}\left(\omega_{n, m} r / c_{2}\right)+\gamma_{n, m} n_{n}\left(\omega_{n, m} r / c_{2}\right), \quad b<r<a
\end{array}\right.
$$

Nowhere in the derivation of the general orthogonality condition in Eqs. (10.3.8) and (10.3.9) was it required that the fluid be homogeneous. It was implicit to the derivation that the mode functions are continuous and that their gradient is continuous. These requirements in the present system are assured by the continuity conditions at $r=b$. The general statements call for an integration over the entire domain, but the modes have different functional dependence in each fluid. Therefore, we decompose the volume integral into integrals over $0<r<b$ and $b<r<a$. Each integral accounts for the different definitions of the mode function in the inner and outer regions. Because the functions are axisymmetric, we may use $d \mathcal{V}=2 \pi(r \sin \Psi)(r d \psi) d r$. The volume of the sphere is $(4 / 3) \pi a^{3}$. Thus, the condition tells us that

$$
\begin{aligned}
\iiint_{\mathcal{V}} \Psi_{n, m} \Psi_{j, \ell} d \mathcal{V}= & 2 \pi \int_{0}^{\pi}\left[\int_{0}^{b} \Psi_{n, m} \Psi_{j, \ell} r^{2} d r+\int_{b}^{a} \Psi_{n, m} \Psi_{j, \ell} r^{2} d r\right] \sin \Psi d \psi \\
= & 2 \pi B_{n, m}^{(1)} B_{j, \ell}^{(1)} \int_{0}^{\pi} P_{n}(\cos \psi) P_{j}(\cos \psi)(\sin \psi) d \psi\left[\int_{0}^{b} F_{n, m}(r) F_{j, \ell}(r) r^{2} d r\right. \\
& \left.+\int_{b}^{a} F_{n, m}(r) F_{j, \ell}(r) r^{2} d r\right]=\delta_{n, j} \delta_{m, \ell}\left(\frac{4}{3} \pi a^{3}\right)
\end{aligned}
$$

If $n$ and $j$ are different, the left side vanishes regardless of $m$ and $\ell$ because the Legendre functions are orthogonal according to Eq.(7.1.16). The condition that is satisfied by the $F$ function is obtained by setting $j=n$. This property reduces the orthogonality integral to

$$
B_{n, m}^{(1)} B_{n, \ell}^{(1)}\left[\int_{0}^{b} F_{n, m}(r) F_{j, \ell}(r) r^{2} d r+\int_{b}^{a} F_{n, m}(r) F_{j, \ell}(r) r^{2} d r\right]=\delta_{m, \ell} \frac{2 n+1}{3} a^{3}
$$

Normalization of $B_{n, m}^{(1)}$ is obtained by setting $\ell=m$.
To close the analysis, let us verify that there is a zero frequency mode. If $\omega=0$, then $k_{1}=k_{2}=0$. In that case, $j_{n}\left(k_{1} r\right), j_{n}\left(k_{2} r\right)$, and $n_{n}\left(k_{2} r\right)$ are identically zero unless $n=0$. Thus, a zero frequency mode can only occur for $n=0$. Furthermore, because $n_{0}(0)$ is infinite, the zero frequency mode requires that $C^{(2)}=0$. Substitution of these values, combined with the property that $j_{0}^{\prime}(0)=0$ and $P_{0}(\cos \Psi) \equiv 1$, leads to recognition that only the second of Eq. (2) is not trivial when $\omega=0$. The second equation in that case leads to $B^{(2)}=B^{(1)}$. Thus, the zero frequency mode is a constant value that is the same in both regions.

$$
\Psi_{0,1}^{(1)}=\Psi_{0,1}^{(2)}=B^{(1)}, \quad 0<r<b
$$

This fits the general property of zero frequency cavity modes.

### 10.4 Approximate Methods

Several factors limit the types of enclosed domains that can be analyzed mathematically. The primary issues pertain to the shape of the cavity and the nature of the surface bounding the enclosed region. There is an obvious need for a formulation that is more general. Discretizations using either boundary elements or finite elements fit this specification. Their implementations for radiation from a vibrating surfaces, which were described in Chap. 7, are readily modified to address an enclosed region. The fundamental alteration of a boundary element formulation entails reversing the surface normal $\bar{n}$ to point into the domain enclosed by the boundary, rather into the region exterior to the boundary. A merit of this formulation is that concern for the issue of "forbidden frequencies" is not relevant. Finite element implementations also benefit when applied to a cavity because there is no need for infinite elements or any of the other techniques for modeling an infinite domain.

In any event, a model created with either discretization method is likely to feature a large number of variables that are the pressures at points an a spatial grid. The fineness of such a grid must increase as the frequency increases. Consequently, the computational resources required to implement numerical simulation techniques makes them prohibitive for the early stages of the design process, as well as for gaining a preliminary insight when an existing system exhibits anomalous behavior.

Two approximate approaches we shall consider generate relatively small models. The first formulation, which originated with Rayleigh, ${ }^{9}$ has two objectives: estimate the lower natural frequencies and find a series solution that best fits the actual cavity modes. The development retains the restriction that the boundary be rigid and/or pressure-release. The second method is known as Dowell's approximation. ${ }^{10}$ It is used to model systems in which the walls are flexible, usually because they are part of an elastic structure.

### 10.4.1 The Rayleigh Ratio and Its Uses

We consider here a limited task: Determine the natural frequencies and modes of an enclosed region whose shape is not amenable to formal mathematical analysis. The concepts we shall explore are applicable to a variety of physical systems. Indeed, they usually are encountered in a course on mechanical and structural vibration. Our implementation is based on the expressions for the acoustic kinetic and potential energy.

## Estimation of the Fundamental Frequency

The foundation for the development is the statement of the basic work-energy principle, Eq. (4.5.30). Because we have restricted our attention to cavities whose walls are rigid (no movement) or pressure-release (no traction force), no power is transferred out of the fluid domain. Hence, the acoustical energy $K E+P E$ must be constant. We know that the cavity modes $\Psi_{n}$ are real functions, and Euler's equation gives $\bar{v}(\bar{x}, t)=\operatorname{Re}\left(i \nabla \Psi_{n} e^{i \omega t}\right) /\left(\rho_{0} \omega\right)$. This means that the pressure everywhere is in-phase, as is the particle velocity, with the particle velocity leading the pressure by $90^{\circ}$. Consequently, at instants separated by $T / 2=\pi / \omega$, the kinetic energy is a maximum and the potential energy is zero. Midway between these instants, the potential energy is maximum and the kinetic energy is zero. It follows from Eq. (4.5.30) that $\max (K E)=\max (P E)$. When we use the field properties of the $n$th mode to form the energy expressions, we find that

$$
\begin{gather*}
\max (K E)=\frac{1}{2} \iiint_{\mathcal{V}} \frac{1}{\rho_{0}\left(\omega_{n}\right)^{2}} \nabla \Psi_{n} \cdot \nabla \Psi_{n} d \mathcal{V} \\
\max (P E)=\frac{1}{2} \iiint_{\mathcal{V}} \frac{1}{\rho_{0} c^{2}}\left(\Psi_{n}\right)^{2} d \mathcal{V} \tag{10.4.41}
\end{gather*}
$$

[^41]Equating these expressions yields

$$
\begin{equation*}
\left(\omega_{n}\right)^{2}=\frac{\iiint_{\mathcal{V}} \frac{1}{\rho_{0}} \nabla \Psi_{n} \cdot \nabla \Psi_{n} d \mathcal{V}}{\iiint_{\mathcal{V}} \frac{1}{\rho_{0} c^{2}}\left(\Psi_{n}\right)^{2} d \mathcal{V}} \tag{10.4.42}
\end{equation*}
$$

This relation is of no direct use, because the analysis leading to a mode function also gives the natural frequency. Its value lies in how Rayleigh generalized it. The numerator and denominator are said to be functionals of $\Psi_{n}$. They shall be denoted as $\mathcal{N}\left[\Psi_{n}(\bar{x})\right]$ and $\mathcal{D}\left[\Psi_{n}(\bar{x})\right]$. The Rayleigh ratio $\mathcal{R}$ is the same ratio of functionals as above, except that $\Psi_{n}$ is replaced by a function $\Psi(\bar{x})$ that is not necessarily a cavity mode,

$$
\begin{equation*}
\mathcal{R}=\frac{\mathcal{N}[\Psi(\bar{x})]}{\mathcal{D}[\Psi(\bar{x})]} \equiv \frac{c^{2} \iiint_{\mathcal{V}} \nabla \Psi \cdot \nabla \Psi d \mathcal{V}}{\iiint_{\mathcal{V}} \Psi^{2} d \mathcal{V}} \tag{10.4.43}
\end{equation*}
$$

The form of this relation is based on the cavity fluid being homogeneous, which is typical for an enclosed region. This restriction allows us to factor $\rho_{0}$ and $c$ out of the integrals. Nevertheless, the concepts that will emerge are equally valid for a heterogeneous medium.

The function $\Psi(\bar{x})$ is said to be a trial function. It should be selected to satisfy two requirements. The first stems from the fact that the pressure within the fluid must be continuous because the particle velocity must be finite. Therefore, $\Psi(\bar{x})$ must be continuous everywhere. It also is necessary that any function we select be zero on any portion of the surface that is pressure-release. The converse condition is that $\Psi$ should not evaluate to zero on the rigid portion of the boundary. (Satisfaction of the latter condition is not mandatory, but it is advisable.) Why do we need to worry about these alternatives? The mathematical explanation lies in the calculus of variations. ${ }^{11}$ A heuristic explanation comes from the observation that the derivation of the Rayleigh ratio required that there be no power transfer at the boundary. We also know that the pressure and normal velocity cannot be specified concurrently anywhere. If $\Psi=0$ on the pressure-release portion of the boundary, the trial function inherently is consistent with zero power flow across that region, regardless of the value of $\bar{n} \cdot \nabla \Psi$. On the other hand, if $\Psi$ is not zero on the rigid portion of the boundary, then the fact that the principle is based on zero power flow will lead to a result that corresponds to taking $\bar{n} \cdot \nabla \Psi$ to be zero there. This argument does not explain why $\Psi=0$, rather than $\bar{n} \cdot \nabla \Psi=0$ is the condition that must be set. Nevertheless, it is an inviolable requirement that the trial function be zero at any point on the surface that is

[^42]pressure-release. (In the calculus of variations, this is said to be a geometric or imposed boundary condition.)

To identify some important properties of $\mathcal{R}$, we represent the trial function in terms of the actual mode functions of the cavity. We do not know these functions, but if we did, such a modal series would feature a set of coefficients $a_{n}$,

$$
\begin{equation*}
\Psi=\sum_{n=1}^{\infty} a_{n} \Psi_{n} \tag{10.4.44}
\end{equation*}
$$

It will be noted that no allowance has been made in this series for the $n=0$ mode, which is the zero frequency mode for an enclosed region whose entire boundary is rigid. We will examine this case after we have completed the derivations.

Although we do not know the mode functions, we know that they satisfy the orthogonality conditions in Eqs. (10.3.8) and (10.3.9). In the latter, $k_{n}$ is the wavenumber at the $n$th natural frequency, $k_{n} \equiv \omega_{n} / c$. We consider the mode function to be normalized in accord with the first of those conditions. (Normalization merely expedites the derivation.) Thus, substitution of the modal series description of $\Psi$ into the energy functionals leads to

$$
\begin{align*}
\mathcal{N}[\Psi(\bar{x})]=c^{2} \sum_{n=1}^{\infty} \sum_{j=1}^{\infty} a_{n} a_{j} \iiint_{\mathcal{V}} \nabla \Psi_{n} \cdot \nabla \Psi_{j} d \mathcal{V} & =c^{2} \sum_{n=1}^{\infty}\left(k_{n}\right)^{2} a_{n}^{2} \mathcal{V}  \tag{10.4.45}\\
\mathcal{D}[\Psi(\bar{x})]=\sum_{n=1}^{\infty} \sum_{j=1}^{\infty} a_{n} a_{j} \iiint_{\mathcal{V}} \Psi_{n} \Psi_{j} d \mathcal{V} & =\sum_{n=1}^{\infty} a_{n}^{2} \mathcal{V}
\end{align*}
$$

The modes have been ordered in ascending order of their natural frequency, so that $k_{n} / k_{1}>1$ if $n>1$. Hence, using the preceding to form $\mathcal{R}$, followed by factoring $\left(k_{1}\right)^{2} \equiv\left(\omega_{1} / c\right)^{2}$ out of the numerator leads to

$$
\begin{equation*}
\mathcal{R}=\frac{\mathcal{N}[\Psi(\bar{x})]}{\mathcal{D}[\Psi(\bar{x})]}=\left(\omega_{1}\right)^{2} \frac{\sum_{n=1}^{\infty}\left(\frac{k_{n}}{k_{1}}\right)^{2} a_{n}^{2}}{\sum_{n=1}^{\infty} a_{n}^{2}} \tag{10.4.46}
\end{equation*}
$$

Except for the $n=1$ term, each term in the numerator's sum is greater the corresponding term in the denominator's sum. It follows that

$$
\begin{equation*}
\mathcal{R} \geq\left(\omega_{1}\right)^{2} \tag{10.4.47}
\end{equation*}
$$

with the equality sign being the case only if all $a_{n}$ other than $a_{1}$ are zero, which means that $\Psi=\Psi_{1}$. In other words, we have shown that the Rayleigh ratio gives an upper bound to the square of the fundamental frequency of the cavity.

This does not tell us the quality of the upper bound-is it too high by $10 \%$, $100 \%$, or much more? To examine this, let us suppose that the $\Psi$ function is close to the first mode multiplied by a scaling factor. To indicate this proximity, let $a_{1}$ in Eq. (10.4.44) be the scaling factor and let all of the other coefficients be small values by setting $a_{n}=\varepsilon a_{n}^{\prime}, n=2,3, \ldots$, where $|\varepsilon| \ll 1$ and all $\left|a_{n}^{\prime}\right| \leq 1$. Equation (10.4.46) correspondingly indicates that

$$
\begin{equation*}
\mathcal{R}=\left(\omega_{1}\right)^{2} \frac{a_{1}^{2}+\varepsilon^{2} \sum_{n=2}^{\infty}\left(\frac{k_{n}}{k_{1}}\right)^{2}\left(a_{n}^{\prime}\right)^{2}}{a_{1}^{2}+\varepsilon^{2} \sum_{n=1}^{\infty}\left(a_{n}^{\prime}\right)^{2}}=\left(\omega_{1}\right)^{2}\left[1+O\left(\varepsilon^{2}\right)\right] \tag{10.4.48}
\end{equation*}
$$

The notation $O\left(\varepsilon^{2}\right)$ refers to a number whose order of magnitude is $\varepsilon^{2}$. The preceding tells us that the relative error of the estimate for $\left(\omega_{1}\right)^{2}$ will be less than the deviation of the trial function from the true fundamental mode, provided that the trial function is a reasonably good estimate, $\varepsilon<1$.

This leads to a common use of the Rayleigh ratio. To obtain an estimate of the fundamental frequency of a cavity, we may fabricate a trial function that is zero at all locations where the surface is pressure-release. Substitution of that function into the definition of $\mathcal{R}$ can be expected to yield an estimated fundamental natural frequency that is better than the degree to which the trial function approximates the first mode function.

There is one condition that must be satisfied if this expectation is to be met. As was stated, the error in the trial function relative to the true fundamental mode must be $O(\varepsilon)$, but we do not know the true mode. If we blindly select trial functions, we could make a terrible choice. For example, suppose we were to select a trial function for which $a_{1}=0.1, a_{2}=1$, and all other $a_{n}=0$. This would lead to $\mathcal{R}=$ $0.9901\left[\left(10^{-4}\right)\left(\omega_{1}\right)^{2}+\left(\omega_{2}\right)^{2}\right]$, which obviously is not close to $\left(\omega_{1}\right)^{2}$.

This condition usually can be avoided if one is cognizant of a fundamental aspect: the fundamental mode of a system usually exhibits the minimum number of nodal lines consistent with any pressure-release conditions. If we select a trial function having this property, then the estimate for $\omega_{1}$ can be expected to be less than $40 \%$ above the true fundamental frequency, corresponding to $\mathcal{R}<2 \omega_{1}$. We will take up this issue in the next example.

When all of the boundaries of the cavity are rigid, the fundamental mode is a uniform pressure and the fundamental frequency is zero. We do not need the Rayleigh ratio to tell us this, but that is what Eq.(10.4.48) indicates will be approximated. Can we use our knowledge of the nature of the zero frequency mode to obtain an approximation of the lowest nonzero natural frequency? The answer lies in the observation that if the trial function is orthogonal to the first mode, then the coefficient $a_{1}$ will be zero. In that case, Eq. (10.4.46) tells us that $\mathcal{R} \geq\left(\omega_{2}\right)^{2}$, and Eq. (10.4.48) tells us that the error in the estimate of $\left(\omega_{2}\right)^{2}$ obtained from $\mathcal{R}$ will be much less than the amount by which the trial function differs from the second mode. Thus, we must modify the trial function in a manner that leads to $a_{1}=0$. The zero frequency
mode is a constant value everywhere. We may scale a mode by any factor, so we may set $\Psi_{0}(\bar{x})=1$. (This is what would be obtained from the standard normalization of modes.) A trial function $\Psi(\bar{x})$ is orthogonal to $\Psi_{0}$ if

$$
\begin{equation*}
\iiint_{\mathcal{V}} \Psi_{0} \Psi d \mathcal{V}=\iiint_{\mathcal{V}}(1) \Psi d \mathcal{V}=0 \tag{10.4.49}
\end{equation*}
$$

This integral is the mean value of $\Psi$ multiplied by the size of the domain.
Thus, if we wish to determine the fundamental nonzero frequency of a rigidwalled cavity, we should employ a trial function whose mean value is zero. If our initial choice for $\Psi$ does not meet this specification, we can obtain a corrected function $\hat{\Psi}(\bar{x})$ by subtracting the mean value of $\Psi(\bar{x})$, that is,

$$
\begin{equation*}
\text { Rigid - walledcavity }: \hat{\Psi}(\bar{x})=\Psi(\bar{x})-\Psi_{\text {mean }}, \quad \Psi_{\text {mean }}=\frac{1}{\mathcal{V}} \iiint_{\mathcal{V}} \Psi(1) d \mathcal{V} \tag{10.4.50}
\end{equation*}
$$

Using $\hat{\Psi}$ to construct the Rayleigh ratio will lead to $R \geq\left(\omega_{2}\right)^{2}$.

EXAMPLE 10.11 Consider a two-dimensional cavity whose sides at $x=0$, $x=L$, and $y=0$ are rigid and whose side at $y=H$ is pressure-release. Identify three alternative trial functions that may be used to estimate the fundamental natural frequency of the system. Then, determine which gives the best estimate when $H=L / 4, H=L / 2$, and $H=2 L$.

## Significance

Because the analytical solution for the modal properties is readily obtained, this example will demonstrate how to interpret the performance of the Rayleigh ratio approximation.

## Solution

The best way to select a trial function usually is to consider a similar system for which the modes are known. However, we will proceed as though we had no prior experience with a two-dimensional cavity in order to understand how an unfamiliar system might be approached. The boundaries correspond to constant values of $x$ or $y$, so we consider trial functions whose form is $\Psi=f(x) g(y)$. If $g(y)$ were chosen to be a constant, satisfaction of the pressure-release condition at $y=H$ would require that the constant be zero. The next simplest function is $g(y)$ being linear. A linear function whose value is zero at $y=H$ is $g(y)=H-y$. A constant nonzero value of $f(x)$ would not violate any requirements because $x=0$ and $x=L$ are rigid, which requires that $\Psi$ not be zero at those locations. The overall magnitude of $f(x)$, as well as $g(y)$, is irrelevant because any scaling factor will cancel when the Rayleigh ratio is formed. We therefore may set $f(x)=1$, which leads to our
first choice for a trial function, $\Psi_{1}=H-y$. In general, monomials, that is, $x$ or $y$ to an integer power, are the easiest to work with because they lead to simple integrals. Because $\Psi_{1}$ may be written as $x^{0}(H-y)$, it seems reasonable to allow for variation in the $x$ direction by raising the power to one, which leads to $\Psi_{2}=x(H-y)$. This will test what happens if the trial function is zero on a surface that is not pressurerelease. For the third function, let us increase the power of the $y$ dependence, so that $\Psi_{3}=x(H-y)^{2}$. A portion of the boundary is pressure-release, so the issue of eliminating the zero frequency mode is irrelevant.

The analytical steps leading to $\mathcal{R}$ for each trial function, based on integration over $0<x<L, 0<y<H$, are tabulated below.

| $\Psi$ | $H-y$ | $x(H-y)$ | $x(H-y)^{2}$ |
| :---: | :---: | :---: | :---: |
| $\nabla \Psi$ | $-\bar{e}_{y}$ | $(H-y) \bar{e}_{x}-x \bar{e}_{y}(H-y)^{2} \bar{e}_{x}-2 x(H-y) \bar{e}_{y}$ |  |
| $\mathcal{N}[\Psi(\bar{x})]$ | $c^{2} H L$ | $c^{2} \frac{H L\left(H^{2}+L^{2}\right)}{3}$ | $c^{2}\left(\frac{H^{5} L}{5}+\frac{4 H^{3} L^{3}}{9}\right)$ |
| $\mathcal{D}[\Psi(\bar{x})]$ | $\frac{H^{3} L}{3}$ | $\frac{H^{3} L^{3}}{9}$ | $\frac{H^{5} L^{3}}{15}$ |
| $\mathcal{R}$ | $c^{2} \frac{3}{H^{2}}$ | $c^{2}\left(\frac{3}{H^{2}}+\frac{3}{L^{2}}\right)$ | $c^{2}\left(\frac{20}{3 H^{2}}+\frac{3}{L^{2}}\right)$ |
| $\left.\mathcal{R}\left(\frac{L}{c}\right)^{2}\right\|_{H=L / 4}$ | $\mathbf{4 8}$ | 51 | 109.7 |
| $\left.\mathcal{R}\left(\frac{L}{c}\right)^{2}\right\|_{H=L / 2}$ | $\mathbf{1 2}$ | 15 | 29.67 |
| $\left.\mathcal{R}\left(\frac{L}{c}\right)^{2}\right\|_{H=2 L}$ | $\mathbf{0 . 7 5}$ | 3.75 | 4.67 |

The value of $\mathcal{R}$ cannot be less than the square of the fundamental natural frequency. The consequence is the general fact that the smallest value of $R$ is closest to the true value of $\omega_{1}^{2}$. The lowest value for each $H / L$ is listed in bold type. In each case, the best estimate is obtained with $\Psi_{1}$. The actual fundamental frequency is either $\omega_{0,1}=0.5 \pi c / H$ or $\omega_{1,0}=\pi c / L$, depending on the size of $H$ relative to $L$. These values and the lowest estimate obtained from the Rayleigh ratio are listed below.

| $H / L$ | $\omega_{0,1}\left(\frac{L}{c}\right)$ | $\omega_{1,0}\left(\frac{L}{c}\right)$ | $\mathcal{R}^{1 / 2}\left(\frac{L}{c}\right)$ |
| :---: | :---: | :---: | :---: |
| 0.25 | 6.283 | $\mathbf{3 . 1 4 2}$ | 6.928 |
| 0.5 | $\mathbf{3 . 1 4 2}$ | $\mathbf{3 . 1 4 2}$ | 3.464 |
| 2 | $\mathbf{0 . 7 8 5}$ | 3.142 | 0.8660 |

The true fundamental frequency is shown in bold. The error of the Rayleigh ratio estimate for $H / L=0.5$ and 2 is $10 \%$, but it is greater than $100 \%$ for $H / L=0.25$.

To understand why the estimate for $H / L=0.25$ is poor, we need to consider the two possible fundamental mode functions, $\Psi_{0,1}=\cos (0.5 \pi y / H)$ and $\Psi_{1,0}=$
$\cos (\pi x / L)$. If $H>L / 2$, the former is the fundamental mode. In it, there is no variation in the $x$ direction and its value decreases monotonically with increasing $y$. The first trial function also does not depend on $x$, and the fact that it increases with increasing $y$, rather than decreasing, is irrelevant. In contrast, when $H<L / 2$, the fundamental mode is $\Psi_{1,0}$, which is independent of $y$ and decreases monotonically with increasing $x$. None of the trial functions matches this independence on $y$. Indeed, the results suggest that they are approximating $\Psi_{0,1}$.

There is no simple rule for selecting a trial function. Picking it to match the fundamental mode of a similar system is a good idea. Another aspect is recognizing that the fundamental mode will entail the minimum fluctuation in all directions consistent with the requirement that it be zero at boundaries that are pressure-release. If it is crucial to obtain a good estimate, multiple trial functions may be considered, as was done here. Ultimately, a better approach is to use several trial functions, which is the procedure developed in the next section.

## Rayleigh-Ritz Method

The previous example closed with a suggestion is that the Rayleigh ratio could be evaluated for several trial functions. Whichever function led to the smallest value of $\mathcal{R}$ would mark the best estimate of $\left(\omega_{1}\right)^{2}$. This is a flawed approach because it does not provide a definitive path to an improved answer. Much better results, including estimates of mode functions, may be obtained from another property of $\mathcal{R}$.

Suppose that the trial function is close to the $j$ th cavity mode, where $j$ may be any integer. Proximity of $\Psi$ to $\Psi_{j}$ does not preclude them differing by a scaling factor. Therefore, this proximity would be manifested in the modal series representation of the trial function, Eq. (10.4.46), by the coefficient $a_{j}$ being much larger than any other coefficient. We indicate this by setting $a_{n}=\varepsilon a_{n}^{\prime}, n \neq j$, where $|\varepsilon| \ll 1$. In other words, the series representation is

$$
\begin{equation*}
\Psi=a_{j} \Psi_{j}+\sum_{\substack{n=1 \\ n \neq j}}^{\infty} \varepsilon a_{n}^{\prime} \Psi_{n} \tag{10.4.51}
\end{equation*}
$$

When $a_{n}$ in Eq. (10.4.46) is replaced by $\varepsilon a_{n}^{\prime}$, the result is

$$
\begin{equation*}
\mathcal{R}=\left(\omega_{j}\right)^{2} \frac{a_{j}^{2}+\varepsilon^{2} \sum_{\substack{n=1 \\ n \neq j}}^{\infty}\left(\frac{k_{n}}{k_{j}}\right)^{2}\left(a_{n}^{\prime}\right)^{2}}{a_{j}^{2}+\varepsilon^{2} \sum_{\substack{n=1 \\ n \neq j}}^{\infty}\left(a_{n}^{\prime}\right)^{2}} \tag{10.4.52}
\end{equation*}
$$

The case where the trial function is exactly proportional to the $j$ th normalized mode corresponds to $\varepsilon=0$. This is important because differentiation of the above expression with respect to $\varepsilon$ leads to

$$
\begin{equation*}
\left.\frac{d \mathcal{R}}{d \varepsilon}\right|_{\varepsilon=0}=0 \tag{10.4.53}
\end{equation*}
$$

To understand this property, let us pretend that we know the cavity modes. Then, we actually could construct the modal series by multiplying any $\Psi_{j}$ by a coefficient $a_{j}$ and then adding to that term small contributions of the other modes, with each contribution scaled by a small parameter $\varepsilon$. This construction would allow us to compute the numerator and denominator of the Rayleigh ratio. Altering the value of $\varepsilon$ would yield a different value of $\mathcal{R}$. In a plot of $\mathcal{R}$ as a function of $\varepsilon$, the slope would be horizontal, that is, $d R / d \varepsilon$ would be zero, at $\varepsilon=0$. This is the property of stationarity.

Of course, we do not know the actual mode function. Instead, the concept is to select a trial function that has adjustable parameters. We can use the stationarity property to find the conditions those parameters must satisfy in order that the trial function matches a mode function. It would be problematic to attempt to pursue this concept by selecting a single trial function of position that contains adjustable parameters. It is far easier to use two or more fully prescribed trial functions of position and to let the adjustable parameters be weighting coefficients for their combination.

To see how this concept may be implemented, suppose we have two identified trial functions, $\Psi_{1}$ and $\Psi_{2}$. Rather than evaluating $\mathcal{R}$ for each function, as we did in the previous example, let us form a new trial function that is a linear combination of $\Psi_{1}$ and $\Psi_{2}$. That is, let

$$
\begin{equation*}
\Psi=b_{1} \Psi_{1}+b_{2} \Psi_{2} \tag{10.4.54}
\end{equation*}
$$

Substitution of this expression into the functionals in Eq.(10.4.43) converts $\mathcal{N}$ and $\mathcal{D}$ and therefore $\mathcal{R}$, to algebraic functions of $b_{1}$ and $b_{2}$. The values of $b_{1}$ and $b_{2}$ that lead to a $\Psi$ function that most closely fits a mode are marked by $\mathcal{R}$ having an extreme value.

To understand how we may use this property, suppose we were to create a threedimensional Cartesian plot in which $b_{1}$ and $b_{2}$ are plotted along two axes and the value of $\mathcal{R}$ as a function of a $b_{1}$ and $b_{2}$ pair is plotted on the third axis. We wish to identify points in the $b_{1} b_{2}$ plane that lead to extreme values of $\mathcal{R}$. We could do so by searching for a high or low point or any point where it levels out locally. However, there is no reason to actually construct such a plot, because the point we seek is characterized by the gradient with respect to $b_{1}$ and $b_{2}$ being zero. The components of this gradient are the respective partial derivatives, so we seek the condition for which

$$
\begin{equation*}
\frac{\partial \mathcal{R}}{\partial b_{1}}=0, \quad \frac{\partial \mathcal{R}}{\partial b_{2}}=0 \tag{10.4.55}
\end{equation*}
$$

This does not mean that satisfaction of these conditions will yield the actual mode function. Rather, it states that if we start off with a two-term series representation of a mode, the coefficients $b_{1}$ and $b_{2}$ that give the closest fit to a mode function are those for which the above derivatives vanish.

We may generalize this notion by considering an $N$-term series, which requires that we identify $N$ suitable trial functions. They must be a linearly independent set,
which means that we cannot obtain one of them as a linear combination of the others. The series is

$$
\begin{equation*}
\Psi=\sum_{n=1}^{N} b_{n} \Psi_{n}(\bar{x}) \tag{10.4.56}
\end{equation*}
$$

Substitution of this expression into the numerator and denominator functionals converts them to quadratic sums,

$$
\begin{gather*}
\mathcal{N}=c^{2} \iiint_{\mathcal{V}}\left(\sum_{n=1}^{N} b_{n} \nabla \Psi_{n}\right) \cdot\left(\sum_{j=1}^{N} b_{j} \nabla \Psi_{j}\right) d \mathcal{V}=\sum_{n=1}^{N} \sum_{j=1}^{N} \mathcal{N}_{j, n} b_{j} b_{N}  \tag{10.4.57}\\
\mathcal{D}=\iint_{\mathcal{V}}\left(\sum_{n=1}^{N} b_{n} \Psi_{n}\right)\left(\sum_{j=1}^{N} b_{j} \Psi_{j}\right) d \mathcal{V}=\sum_{n=1}^{N} \sum_{j=1}^{M} \mathcal{D}_{j, n} b_{j} b_{n}
\end{gather*}
$$

where the coefficients are

$$
\begin{gather*}
\mathcal{N}_{j, n}=c^{2} \iiint_{\mathcal{V}} \nabla \Psi_{j} \cdot \nabla \Psi_{n} d \mathcal{V} \\
\mathcal{D}_{j, n}=\iiint_{\mathcal{V}} \Psi_{j} \Psi_{n} d \mathcal{V} \tag{10.4.58}
\end{gather*}
$$

This reduces $\mathcal{R}$ to a function of the $b_{m}$ coefficients, so the stationarity property requires that

$$
\begin{equation*}
\frac{\partial \mathcal{R}}{\partial b_{m}}=0, \quad m=1,2, \ldots, M \tag{10.4.59}
\end{equation*}
$$

According to Eq.(10.4.43), $\mathcal{R}=\mathcal{N} / \mathcal{D}$, so we have

$$
\begin{equation*}
\frac{\partial \mathcal{R}}{\partial b_{m}}=\frac{1}{\mathcal{D}^{2}}\left(\mathcal{D} \frac{\partial \mathcal{N}}{\partial b_{m}}-\mathcal{N} \frac{\partial \mathcal{D}}{\partial b_{m}}\right) \equiv \frac{1}{\mathcal{D}}\left(\frac{\partial \mathcal{N}}{\partial b_{m}}-\mathcal{R} \frac{\partial \mathcal{D}}{\partial b_{m}}\right)=0 \tag{10.4.60}
\end{equation*}
$$

Because $\mathcal{D}$ is finite, it must be that

$$
\begin{equation*}
\frac{\partial \mathcal{N}}{\partial b_{m}}-\mathcal{R} \frac{\partial \mathcal{D}}{\partial b_{m}}=0, m=1,2, \ldots, M \tag{10.4.61}
\end{equation*}
$$

These equations are known as the Rayleigh-Ritz method. Walther Ritz derived the formulation close to the time of Rayleigh's derivation, but he did so by following a very different route. ${ }^{12}$

[^43]The Rayleigh-Ritz equation, Eq.(10.4.61), is quite general. To derive its specific form, we differentiate Eq. (10.4.57) with respect to a specific $b_{m}$. This operation is

$$
\begin{align*}
\frac{\partial \mathcal{N}}{\partial b_{m}}= & \sum_{n=1}^{N} \sum_{j=1}^{N} \mathcal{N}_{j, n} \frac{\partial}{\partial b_{m}}\left(b_{j} b_{n}\right)=\sum_{n=1}^{N} \sum_{j=1}^{N} \mathcal{N}_{j, n}\left(\delta_{j, m} b_{n}+b_{j} \delta_{n, m}\right) \\
& =\sum_{n=1}^{N} \mathcal{N}_{m, n} b_{n}+\sum_{j=1}^{N} \mathcal{N}_{j, m} b_{j}=2 \sum_{n=1}^{N} \mathcal{N}_{m, n} b_{n} \tag{10.4.62}
\end{align*}
$$

The last form results from the symmetry of the coefficients, $\mathcal{N}_{j, m} \equiv \mathcal{N}_{m, j}$, which follows from their definition. A similar result applies to the derivative of $\mathcal{D}$. Thus, we obtain the Rayleigh-Ritz equations,

$$
\begin{equation*}
\sum_{n=1}^{N} \mathcal{N}_{m, n} b_{n}-\mathcal{R} \sum_{n=1}^{N} \mathcal{D}_{m, n} b_{n}=0, \quad m=1,2, \ldots, N \tag{10.4.63}
\end{equation*}
$$

The nature of the equations to solve becomes evident from their matrix form,

$$
\begin{equation*}
[[\mathcal{N}]-\mathcal{R}[\mathcal{D}]]\{b\}=\{0\} \tag{10.4.64}
\end{equation*}
$$

This constitutes a general (matrix) eigenvalue problem, whose solution will yield $M$ eigenvalues $\mathcal{R}_{m} / c$ and corresponding eigenvectors $\{b\}_{m}$.

The eigensolutions have several remarkable properties. Identification of those properties requires considerable analysis, primarily in linear algebra, ${ }^{13}$ so we will only discuss the highlights here. The first is an extension of the upper bound theorem. It asserts that the $r$ th eigenvalue, $\mathcal{R}_{r}$, is an upper bound for the square of the $r$ th natural frequency,

$$
\begin{equation*}
\mathcal{R}_{r} \geq\left(\omega_{r}\right)^{2} \tag{10.4.65}
\end{equation*}
$$

This provides no guidelines as to how much larger is the value of $\mathcal{R}_{m}$. However, unless one makes special effort to find unusual functions, all eigenvalues except the largest can be expected to be within $10 \%$ or less of the respective true values, with the error increasing with increasing mode number.

[^44]The upper bound property is applicable for any set of trial functions. A much stronger property holds if the series of trial functions is formed in a certain way. Let us denote the basis functions of the $N$-term series in Eq. (10.4.56) as $\Psi_{n}^{(N)}$. To this, we add another term consisting of a different trial function $F$, so that the enlarged set of functions is

$$
\begin{equation*}
\Psi_{n}^{(N+1)}=\Psi_{n}^{(N)}, \quad n=1, \ldots, N, \quad \Psi_{N+1}^{(N+1)}=F \tag{10.4.66}
\end{equation*}
$$

Correspondingly, the eigenvalues obtained from the two series are denoted as $R_{r}^{(N)}$, $r=1,2, \ldots, N$, and $R_{r}^{(N+1)}, r=1,2, \ldots, N+1$. The result will satisfy the separation theorem, which states that the eigenvalues for the shorter length series fall between the eigenvalues for the longer series, according to

$$
\begin{equation*}
R_{1}^{(N+1)} \leq R_{1}^{(N)} \leq R_{2}^{(N+1)} \leq R_{2}^{(N)} \leq \ldots \leq R_{N}^{(N)} \leq R_{N+1}^{(N+1)} \tag{10.4.67}
\end{equation*}
$$

In words, this property tells us that as we lengthen the series, the estimate for any natural frequency will decrease or be unchanged. This is a remarkable result in light of the upper bound theorem, because the estimated value must be greater than the true value. Therefore, it must be that lengthening the series will cause the eigenvalues to converge from above to the true natural frequencies.

A corollary of convergence of the eigenvalues is convergence to the true mode functions of the trial function series. This attribute follows from Eq. (10.4.52), which states that if a value of $\mathcal{R}$ is close to $\left(\omega_{j}\right)^{2}$, then it must be that the corresponding $a_{j}$ coefficient is much larger than any of the other coefficients. To construct a mode, let $\left(b_{n}\right)_{r}$ be the $n$th element of the eigenvector corresponding to $\mathcal{R}_{r}$. The trial function series corresponding to the $r$ th eigenvector is

$$
\Psi_{r}^{(N)}=\sum_{n=1}^{N}\left(b_{n}\right)_{r} \Psi_{n}(\bar{x}) \equiv\left[\begin{array}{llll}
\Psi_{1}(\bar{x}) & \Psi_{2}(\bar{x}) & \ldots & \Psi_{N}(\bar{x}) \tag{10.4.68}
\end{array}\right]\{b\}_{r}
$$

The superscript " $N$ " has been inserted to remind us that we have constructed an $N$-term series approximation of the actual cavity mode $\Psi_{r}$.

EXAMPLE 10.12 It is desired to identify the natural frequencies and mode functions of a two-dimensional cavity in the shape of an ellipse defined as $(x / a)^{2}+(y / b)^{2}=1$, where $a$ and $b$ are the semimajor and semiminor axis lengths. The elliptical surface of the cavity is rigid. The mode functions of interest are those that are symmetric with respect to the $x z$ and $y z$ bisecting planes, with a nonzero natural frequency. For the case where $a=2 b$, use series containing two to six terms to determine the natural frequencies and mode functions by the Rayleigh-Ritz method. Assess the frequencies in light of the separation theorem, and compare the results for the first three mode functions.

## Significance

Analysis of the mode functions for an elliptical shape is an imposing problem if it is done by seeking eigensolutions of the Helmholtz equation. Thus, it is systems like this for which the Rayleigh-Ritz method is especially useful. The solution will demonstrate how exploiting a system's symmetry properties can improve the analysis without requiring greater computational effort. The results will exemplify the general performance characteristics of the Rayleigh-Ritz method.

## Solution

The elliptical wall is rigid, so the trial functions should not vanish on the boundary. Using a power series in each direction is acceptable, and those functions are easy to use. One consideration is the zero frequency mode, which this system possesses. When the Rayleigh ratio is used to estimate the fundamental frequency for such systems, the mean value of the basis function is subtracted from the function in order to estimate the fundamental nonzero frequency. We could do the same here by subtracting from each trial function its mean value over the elliptical cross section. In the Rayleigh-Ritz method, many modes are obtained, so a simpler alternative is to include a trial function that is uniform over the cross section. Suppose we use power series for both the $x$ and $y$ dependencies. We form $\Psi$ as a product of the individual series, and designate the independent coefficients as $c_{m}$,

$$
\begin{align*}
\Psi & =\left(\alpha_{0}+\alpha_{1} x+\alpha_{2} x^{2}+\cdots\right)\left(\beta_{0}+\beta_{1} y+\beta_{2} y^{2}+\cdots\right)  \tag{1}\\
& =c_{1}+c_{2} x+c_{3} y+c_{4} x^{2}+c_{5} y^{2}+c_{6} x y+\cdots
\end{align*}
$$

It makes sense to halt the series at a stage where all combinations of $x$ and $y$ up to the same degree have been included. Equation (1) corresponds to truncation at quadratic terms.

Equation (1) would be a valid starting point for a general analysis. However, it has been stipulated that we should find the symmetric mode functions. A function is symmetric with respect to the $x z$ plane if $f(x, y)=f(x,-y)$, and symmetry with respect to the $y z$ plane requires $f(x, y)=f(-x, y)$. These properties will be obtained if the $x$ and $y$ polynomials in Eq. (1) are even functions of $x$ and $y$, which means that they should be polynomials that contain only even powers. Thus, the series we shall use is formed as

$$
\begin{align*}
\Psi & =\left(\alpha_{0}+\alpha_{2} x^{2}+\alpha_{4} x^{4}+\cdots\right)\left(\beta_{0}+\beta_{2} y^{2}+\beta_{4} y^{4}+\cdots\right) \\
& =c_{1}+c_{2} x^{2}+c_{3} y^{2}+c_{4} x^{4}+c_{5} y^{4}+c_{6} x^{2} y^{2}+\cdots \tag{2}
\end{align*}
$$

Note that this series has the same number of terms as Eq.(1), but it is quartic and therefore can be expected to better fit the true modes. The trial functions are the coefficients of the respective $c_{m}$ in Eq. (2). It is useful to work with their dimensionless equivalents, which are

$$
\begin{equation*}
\Psi_{1}=1, \Psi_{2}=\frac{x^{2}}{a^{2}}, \Psi_{3}=\frac{y^{2}}{b^{2}}, \Psi_{4}=\frac{x^{4}}{a^{4}}, \Psi_{5}=\frac{y^{4}}{b^{4}}, \quad \Psi_{6}=\frac{x^{2} y^{2}}{a^{2} b^{2}} \tag{3}
\end{equation*}
$$

Because the functions are even with respect to $x$ and $y$, we may implement Eq. (10.4.58) by integrating over one quadrant and then multiplying the result by four. Hence, the domain of integration will be $0<y<b\left(1-(x / a)^{2}\right)^{1 / 2}, 0<x<a$. These functions have the generic form $\Psi_{j}=(x / a)^{r_{j}}(y / b)^{s_{j}}$. Thus, the coefficients for the numerator of $\mathcal{R}$ are described by

$$
\begin{align*}
\mathcal{N}_{j, m}= & 4 c^{2} \int_{0}^{a} \int_{0}^{b\left(1-(x / a)^{2}\right)^{1 / 2}}\left[\frac{r_{j}}{a}\left(\frac{x}{a}\right)^{r_{j}-1}\left(\frac{y_{j}}{b}\right)^{s_{j}} \bar{e}_{x}+\frac{s_{j}}{b}\left(\frac{x}{a}\right)^{r_{j}}\left(\frac{y}{b}\right)^{s_{j}-1} \bar{e}_{y}\right] \\
= & \left.4 a b c^{2} \int_{0}^{1} \int_{0}^{\left(1-\left(x^{\prime}\right)^{2}\right)^{1 / 2}}\left\{\frac{r_{m}}{a}\left(\frac{x}{a}\right)^{r_{m}-1}\left(\frac{y}{b}\right)^{s_{m}} \bar{e}_{x}+\frac{s_{m}}{b}\left(\frac{x}{a}\right)^{r_{m}}\left(\frac{y}{b}\right)^{s_{m}-1}\right)^{r_{j}+r_{m}-2}\left(y^{\prime}\right)^{s_{j}+s_{m}}\right] d y d x \\
& \left.+\frac{s_{j} s_{m}}{b^{2}}\left(x^{\prime}\right)^{r_{j}+r_{m}}\left(y^{\prime}\right)^{s_{j} s_{m}-2}\right\} d y^{\prime} d x^{\prime} \\
= & 4 c^{2} \int_{0}^{1}\left\{\left(\frac{b}{a}\right) \frac{r_{j} r_{m}}{\left(s_{j}+s_{m}+1\right)}\left(x^{\prime}\right)^{r_{j}+r_{m}-2}\left[1-\left(x^{\prime}\right)^{2}\right]^{\left(s_{j}+s_{m}+1\right) / 2}\right. \\
& \left.+\left(\frac{a}{b}\right) \frac{s_{j} s_{m}}{\left(s_{j}+s_{m}-1\right)}\left(x^{\prime}\right)^{r_{j}+r_{m}}\left[1-\left(x^{\prime}\right)^{2}\right]^{\left(s_{j}+s_{m}-1\right) / 2}\right\} d x^{\prime}, \quad s_{j}>0, s_{m}>0
\end{align*}
$$

Cases where either $s_{j}=0$ or $s_{m}=0$ are readily handled separately. The denominator coefficients are

$$
\begin{align*}
\mathcal{D}_{j, m} & =4 \int_{0}^{a} \int_{0}^{b\left(1-(x / a)^{2}\right)^{1 / 2}}\left(\frac{x}{a}\right)^{r_{j}}\left(\frac{y}{b}\right)^{s_{j}}\left(\frac{x}{a}\right)^{r_{m}}\left(\frac{y}{b}\right)^{s_{m}} d y d x \\
& =4 a b \int_{0}^{1} \frac{1}{\left(s_{j}+s_{m}+1\right)}\left(x^{\prime}\right)^{r_{j}+r_{m}}\left[1-\left(x^{\prime}\right)^{2}\right]^{\left(s_{j}+s_{m}+1\right) / 2} d x^{\prime} \tag{5}
\end{align*}
$$

The most expedient approach is to use numerical methods to evaluate the integrals. A small saving is that $[\mathcal{D}]$ and $[N]$ are symmetric, so it is sufficient to compute only $\mathcal{D}_{j, m}$ and $\mathcal{N}_{j, m}$ for $m \leq j$.

To address the fact that only the value of $b / a$ is specified, let us place the eigenvalue problem into nondimensional form. To that end, let $[\mathcal{N}]=c^{2}[\hat{\mathcal{N}}]$ and $[\mathcal{D}]=a^{2}[\hat{\mathcal{D}}]$. Doing so converts Eq. (10.4.64) to

$$
\begin{equation*}
\left[[\hat{\mathcal{N}}]-\frac{a^{2}}{c^{2}} \mathcal{R}[\hat{\mathcal{D}}]\right]\{b\}=\{0\} \tag{6}
\end{equation*}
$$

The eigenvalues are $\left(a^{2} / c^{2}\right) \mathcal{R}_{n}$, which are approximations of the square of nondimensional natural frequencies $k_{n} a$.

The computer algorithm that was implemented began by defining row vectors for the exponents in Eq. (3) sequenced in the order of the trial functions, such that

$$
[r]=\left[\begin{array}{llllll}
0 & 2 & 0 & 4 & 0 & 2
\end{array}\right], \quad[s]=\left[\begin{array}{llllll}
0 & 0 & 2 & 0 & 4 & 2
\end{array}\right]
$$

The coefficients in Eqs. (4) and (5) were determined numerically. The MATLAB numerical integration routine that was used is quadl. It requires that the function that is integrated depends only on the integration variable. This is achieved by defining $[r]$ and $[s]$ prior to an outer loop over $j=1,2, . ., 6$ and an inner loop over $m=1, \ldots, j$. Inside these loop, anonymous functions defining each integrand are defined with the $r$ and $s$ values referred to explicitly, for example

```
N_integrand=@(x) 4*b_a*r(j)*r(m)*x.^(r(j)+r(m)-2).*(sqrt(1-x.^2))...
    \(.^{\wedge}(s(j)+s(m)+1) /(s(j)+s(m)+1) ;\)
```

The name of each function is the first argument for the quadl routine. Completion of the loops over $m$ and $j$ fills in $[\hat{\mathcal{N}}]$ and $[\hat{\mathcal{D}}]$. The next set of operations is solution of the eigenvalue problem for series lengths ranging from two to six. This is done by creating a loop for $N$ from to six. Within it, submatrices of $[\hat{\mathcal{N}}]$ and $[\hat{\mathcal{D}}]$ are input to the eigenvalue solver. In MATLAB, the operation is [b_vals, eigvals] = eig(N_coeff(1 : N, 1 : N), D_coeff(1 : N, 1 : N));. Each column of b_vals is an eigenvector, and the diagonal elements of eigvals are the corresponding values of $\left(a^{2} / c^{2}\right) \mathcal{R}_{n}$. The diagonal elements of eigvals for each $N$ are saved in a matrix for examination upon completion of all operations.

The eigenvectors are needed to generate the mode functions. A useful way in which they may be visualized is with a surface plot, but doing so entails a change in how Eq. (10.4.68) is evaluated. Suppose $[X]$ and $[Y]$ hold the values of $x / a$ and $y / b$ at a grid of points in the ellipse. (These matrices may be created by the meshgrid command in MATLAB.) Let us expand the summation form of the aforementioned equation, such that

$$
\begin{equation*}
\Psi_{r}^{(N)}=\Psi_{1}(\bar{x})\left(b_{1}\right)_{r}+\Psi_{2}(\bar{x})\left(b_{2}\right)_{r}+\cdots \tag{10.4.69}
\end{equation*}
$$

where $\Psi_{r}^{(N)}$ is the $r$ th mode function, the $\Psi_{j}(\bar{x})$ are trial functions evaluated at $\bar{x}$, and $\left(b_{j}\right)_{r}$ is the $j$ th element in column $r$ of the eigenvector matrix. (In the present case, this matrix is b_vals.) If we define $\Psi_{j}(\bar{x})$ to be a matrix evaluated at each point in the grid, then the result will be the value of $\Psi_{r}^{(N)}$ at all grid points. The sum may be evaluated cumulatively. MATLAB code for these operations is Psi_mode_r=zeros(size(X)); for $j=1: N$; Psi_j= X.^r(j)* (Y/b_a).^s(j);. Psi_mode_r=Psi_mode_r+Psi_j*b_vals(j,r); end;

Because $\Psi_{1}=1$, so that $\nabla \Psi_{1}=0$, the first row and column of $[\mathcal{N}]$ are zero, which leads to $\mathcal{R}_{1}=0$ for any series length. The corresponding mode function is a constant. The results for the values of $\mathcal{R}_{n}(a / c)^{2}$ are described in the tabulation. The separation theorem states that any value in a column for $N>2$ should be less than or equal to the value to the left of it, and greater than or equal to the value above it. The tabulated values are consistent with this theorem. We also see that the last value of $\mathcal{R}$ at any series length is the one that changes most when a term is added to the series. This means that the last mode, whose natural frequency is highest, will be the most in error. The typical overall trend is that at any series length, the error
in the eigenvalue, which may be identified by comparison to a solution of the field equations or by examining the trend in a tabulation, increases with increasing mode number. All except the highest may be expected to be reasonably close to the true value.

|  | $N=2$ | $N=3$ | $N=4$ | $N=5$ | $N=6$ |
| :--- | :---: | :---: | :---: | :---: | :---: |
| $\mathcal{R}_{1}(a / c)^{2}$ | 0 | 0 | 0 | 0 | 0 |
| $\mathcal{R}_{2}(a / c)^{2}$ | 16.0000 | 15.4534 | 11.9353 | 11.8629 | 11.7446 |
| $\mathcal{R}_{3}(a / c)^{2}$ |  | 74.5466 | 70.7570 | 53.9301 | 46.5883 |
| $\mathcal{R}_{4}(a / c)^{2}$ |  |  | 87.3552 | 86.8430 | 69.9076 |
| $\mathcal{R}_{5}(a / c)^{2}$ |  |  |  | 354.1028 | 177.7429 |
| $\mathcal{R}_{6}(a / c)^{2}$ |  |  |  |  |  |
|  |  |  |  |  |  |

The mode functions converge more slowly than the natural frequencies. The changes as the series is lengthened are less readily quantified, so comparisons usually are visual. Figure 1 compares the result of using Eq. (10.4.68) to evaluate $\Psi_{2}$ for several series lengths. The differences are not great. Indeed, the high quality of the $N=2$ result, for which the corresponding value of $\mathcal{R}_{2}$ is $36 \%$ too large, is somewhat surprising.


Figure 1.

Length limitations prevent a full comparison of the mode functions for various series lengths. Figure 2 shows some modes for $N=6$. There seems to be little evidence of the modal distributions for a circular cross section, but $\Psi_{2}$ does resemble the $(1,0)$ mode for a rectangular shape whose aspect ratio is $2: 1$.


Figure 2.

### 10.4.2 Dowell's Approximation

Evaluation of the time-domain response within a cavity whose walls are compliant is a core problem for the subject of structural acoustics. Usually, such walls are the surface of an elastic structure. The two media interact in two ways. The pressure exerted by the fluid constitutes a force system that is applied to the structure. Its effect must be incorporated into the structure's equations of motion. The excitation applied to the fluid stems from the requirement that the normal component of its particle velocity and that of the structure be equal everywhere along the surface at which they meet. Analytical solutions of such problems can be found only for very simple circumstances. Numerical modeling techniques often are used, but as noted previously, they might be unwieldy for some applications.

The fundamental formulation of the Rayleigh ratio is valid if the structure is elastic because all that was assumed is that the system is time-invariant and that the sum of kinetic energy and potential energy is conserved in a free vibration. The difficulty with implementation of either the Rayleigh ratio or Rayleigh-Ritz method lies in the selection of trial functions. In addition to selection of one or more functions for the fluid, it would be necessary to formulate a series to describe the structure's displacement. This is where a dilemma arises, because there is no convenient way to assure that the pressure and displacement series satisfy continuity conditions at the fluid-structure interface.

The modeling formulation derived from Dowell's approximation addresses this issue. The analytical method is general. It has the beneficial feature of leading to model equations that feature a relatively small number of variables. However, these benefits are obtained by approximating one of the fundamental fluid-structure interaction mechanisms.

The portion of the fluid's boundary that is the interface with the structure is denoted as $\mathcal{S}_{e}$. The displacement of the structure normal to $\mathcal{S}_{e}$ is an unknown function $w\left(\bar{x}_{s}, t\right)$ that must be determined as part of the analysis. The remainder of the bounding surface is denoted as $\mathcal{S}_{v}$ on which the normal velocity is taken to be a known function $\hat{v}\left(\bar{x}_{s}, t\right)$. Thus, $\mathcal{S}_{v}$ includes any regions that are rigid. Both $w$ and $\hat{v}$ are taken to be positive if they correspond to movement into $\mathcal{V}$, which is the sense in which the surface normal $\bar{n}\left(\bar{x}_{s}\right)$ is defined. Hence, continuity at the surface requires that

$$
\begin{gather*}
\bar{n} \cdot \bar{v}=\dot{w}, \quad \bar{x}_{s} \in \mathcal{S}_{e}  \tag{10.4.70}\\
\bar{n} \cdot \bar{v}=\hat{v}, \bar{x}_{s} \in \mathcal{S}_{v}
\end{gather*}
$$

The core concept of Dowell's approximation is a series representation of pressure in which the basis functions are the cavity modes for a domain having the same shape, but whose walls are rigid. Analysis of such modes has already been discussed, but their actual determination might represent a significant task if the cavity is irregular in shape. We shall proceed on the assumption that we have found those modes, which we denote as $\Psi_{R, n}$, where $n$ is the mode number and " $R$ " serves to remind us that these are rigid-cavity mode functions. The objective is to derive a set of equations whose solution will describe the pressure field within the cavity and the response of the structure as functions of time. Correspondingly, the coefficients of the cavity mode series are taken to be functions of time to be determined. Thus, the ansatz is

$$
\begin{equation*}
p(\bar{x}, t)=\sum_{n=0}^{N} \Psi_{R, n}(\bar{x}) p_{n}(t) \tag{10.4.71}
\end{equation*}
$$

where the summation index $n$ begins from zero as a reminder to include the zero frequency mode in which the pressure is uniform throughout the cavity.

The fundamental nature of Dowell's approximation is embedded in the above series. In the time domain, Euler's equation states that the particle acceleration normal to the surface is proportional to the normal component of $\nabla p$. By definition, $\bar{n}$.
$\nabla \Psi_{R, n}=0$ on the boundary. Hence, the series construction gives a zero acceleration on the surface, whereas the time derivative of Eq. (10.4.70) requires that it be $\ddot{w}$. The development proceeds by ignoring this paradox, but the surface continuity conditions will enter the analysis at a later juncture.

In addition to excitation by movement of the boundary, there might be a set of sources at several locations $\bar{x}_{m}$. Thus, the pressure $p$ is governed by the inhomogeneous wave equation, Eq.(10.3.11), which is

$$
\begin{equation*}
\nabla^{2} p-\frac{1}{c^{2}} \frac{\partial^{2} p}{\partial t^{2}}=-\rho_{0} \sum_{m} \dot{Q}_{m}(t) \delta\left(\bar{x}-\bar{x}_{m}\right) \tag{10.4.72}
\end{equation*}
$$

Equations (10.4.70) are the boundary conditions governing $p(\bar{x}, t)$. The equations satisfied by the rigid-cavity modes are

$$
\begin{gather*}
\nabla^{2} \Psi_{R, n}+\left(k_{r, n}\right)^{2} \Psi_{R, n}=0, \quad \bar{x} \in \mathcal{V}  \tag{10.4.73}\\
\bar{n} \cdot \nabla \Psi_{R, n}=0, \quad \bar{x} \in \mathcal{S}_{e} \cup \mathcal{S}_{v}
\end{gather*}
$$

With the aim of employing Green's theorem, we multiply the field equation for a specific rigid body $\Psi_{R, n}$ mode by $p$, and the field equation for $p$ by $\Psi_{R, n}$. Then, we subtract the latter from the former and integrate over the fluid domain $\mathcal{V}$. The result of these operations is

$$
\begin{align*}
\iiint_{\mathcal{V}} & \left\{p\left[\nabla^{2} \Psi_{R, n}+\left(k_{r, n}\right)^{2} \Psi_{r, n}\right]-\Psi_{r, n}\left[\nabla^{2} p-\frac{1}{c^{2}} \frac{\partial^{2} p}{\partial t^{2}}\right]\right\} d \mathcal{V} \\
& =\rho_{0} \sum_{m} \dot{Q}_{m}(t) \Psi_{R, n}\left(\bar{x}_{m}\right) \tag{10.4.74}
\end{align*}
$$

Green's theorem converts the terms containing the Laplacian to a surface integral according to

$$
\begin{align*}
\iiint_{\mathcal{V}}\left(p \nabla^{2} \Psi_{R, n}-\Psi_{R, n} \nabla^{2} p\right) d \mathcal{V} & \equiv \iiint_{\mathcal{V}}\left[\nabla \cdot\left(p \nabla \Psi_{R, n}-\Psi_{R, n} \nabla p\right)\right] d \mathcal{V} \\
& =-\iint_{\mathcal{S}}\left[p\left(\bar{n} \cdot \nabla \Psi_{R, n}\right)-\Psi_{R, n}(\bar{n} \cdot \nabla p)\right] d \mathcal{S} \tag{10.4.75}
\end{align*}
$$

The minus sign preceding the surface integral stems from the normal direction $\bar{n}$ being defined to point into $\mathcal{V}$.

The surface integral extends over the entire boundary $\mathcal{S}$. The normal derivative of $\Psi_{R, n}$ is zero everywhere on this surface, whereas $\bar{n} \cdot \nabla p$ is related to the surface normal acceleration by Euler's equation and the time derivative of Eq.(10.4.70). Thus, we have established that

$$
\begin{equation*}
\iiint_{\mathcal{V}}\left(p \nabla^{2} \Psi_{R, n}-\Psi_{R, n} \nabla^{2} p\right) d \mathcal{V}=-\iint_{\mathcal{S}_{e}} \Psi_{R, n} \rho_{0} \ddot{w} d \mathcal{S}-\iint_{\mathcal{S}_{v}} \Psi_{R, n} \rho_{0} \frac{\partial \hat{v}}{\partial t} d \mathcal{S} \tag{10.4.76}
\end{equation*}
$$

It is important to recognized that although the rigid-cavity series cannot give the correct particle velocity at the boundary, the implementation of Green's theorem uses the actual boundary velocity. Thus, the formulation accounts in an average sense for the way movement at the boundary influences the pressure.

Substitution of the preceding relation into Eq. (10.4.74) leads to

$$
\begin{align*}
& \iiint_{\mathcal{V}} \Psi_{R, n}\left[\frac{1}{c^{2}} \frac{\partial^{2} p}{\partial t^{2}}+\left(k_{r, n}\right)^{2} p\right] d \mathcal{V}-\iint_{\mathcal{S}_{e}} \Psi_{R, n} \rho_{0} \ddot{w} d S-\iint_{\mathcal{S}_{v}} \Psi_{R, n} \rho_{0} \frac{\partial \hat{v}}{\partial t} d S \\
& =\rho_{0} \sum_{m} \dot{Q}_{m}(t) \Psi_{R, n}\left(\bar{x}_{m}\right) \tag{10.4.77}
\end{align*}
$$

The integrands of the surface integrals consist of known quantities, so they belong on the right side of the equation. Furthermore, we have already represented $p$ as a series of rigid-cavity modes according to Eq.(10.4.71). When we substitute that description of $p$ into the preceding, $\Psi_{R, n}$ multiplies every $\Psi_{R, j}$ in the summation. The first orthogonality condition for cavity modes is

$$
\begin{equation*}
\iiint_{\mathcal{V}} \Psi_{R, n} \Psi_{R, j} d \mathcal{V}=\mathcal{V} \tag{10.4.78}
\end{equation*}
$$

Thus, orthogonality filters out of the summation all modes except number $n$. We thereby obtain a set of uncoupled differential equations for the modal coefficients,

$$
\begin{align*}
\mathcal{V}\left[\ddot{p}_{n}+\left(\omega_{R, n}\right)^{2} p_{n}\right]= & \rho_{0} c^{2} \iint_{\mathcal{S}_{e}} \Psi_{R, n} \ddot{w} d \mathcal{S}+\rho_{0} c^{2} \iint_{\mathcal{S}_{v}} \Psi_{R, n} \frac{\partial \hat{v}}{\partial t} d \mathcal{S}  \tag{10.4.79}\\
& +\sum_{m} \rho_{0} \dot{Q}_{m}(t) \Psi_{R, n}\left(\bar{x}_{m}\right), \quad n=0,1, \ldots, N
\end{align*}
$$

where $\omega_{R, n} \equiv c k_{R, n}$ are the natural frequencies of the rigid cavity. Each of the terms in this expression represents a different physical process: $\ddot{p}_{n}$ represents compressibility, $\left(\omega_{n}\right)^{2} p_{n}$ represents inertia, $\dot{Q}_{m}$ is an excitation that results from injecting mass into the domain, $\partial \hat{v} / \partial \hat{v}$ is a surface source distribution resulting from causing the boundary to move, and $\ddot{w}$ is a surface source distribution resulting from movement of the elastic structure at its interface with the fluid.

Equation (10.4.79) constitutes a set of $N+1$ ordinary differential equations for the $N+1$ pressure coefficients. If there were no structure, these equations would suffice. Our interest here is the interaction of the acoustic fluid and the structure. In such situations, the surface displacement $w$ is a quantity that must be determined. It is here that the equations of motion for the structure enter the analysis. The displacement
of the structure may be described in terms of a set of variables called generalized coordinates and denoted as $q_{j}$. These variables might be the coefficients of a series, or they might be the mesh displacements in a finite element model. Our interest is in systems that are linear and time-invariant, which means that a reference configuration exists in which the system may remain in a state of static equilibrium. In such cases, it is possible to define the generalized coordinates such that zero values for all place the structure at its reference position. For a given system, there are several methods for formulating the differential equations of motion governing the generalized coordinates. Usually, the derivation of those equations is founded on Lagrange's equations, Hamilton's principle, or the principle of virtual work, which are closely related. ${ }^{14}$ The result will be the structure's equations of motion in a canonical form, whose matrix description is

$$
\begin{equation*}
[M]\{\ddot{q}\}+[K]\{q\}=\{F\}+\left\{F_{\text {pressure }}\right\} \tag{10.4.80}
\end{equation*}
$$

The matrices [ $M$ ] and [ $K$ ], which are constant, square, and symmetric, are the inertia and stiffness. Another structural effect is dissipation, but the actual mechanisms by which energy is lost usually are uncertain. Consequently, dissipation is usually introduced at a later stage.

The elements of the column vectors $\{F\}$ and $\left\{F_{\text {pressure }}\right\}$ are generalized forces. The former represents the excitation forces acting on the structure, whereas the latter contain the effects of the surface pressure. Determination of $\{F\}$ is part of the standard process by which equations of motion are derived. Determination of $\left\{F_{\text {pressure }}\right\}$ is an extension of that process. Identification of the force terms begins with a general description of the surface displacement $w\left(\bar{x}_{s}, t\right)$. A linear structural model leads to a description of the displacement of any point as a series in which each generalized coordinate is multiplied by a spatial function. These functions depend on how the structure is modeled, but the details are not important here. The surface normal displacement is obtained by evaluating that series at an arbitrary surface point $\bar{x}_{s}$ and then taking a dot product of that vector with the outward normal $\bar{n}\left(\bar{x}_{s}\right)$. The consequence is that $w$ also is represented as a linear series in which the spatial functions depend on $\bar{x}_{s}$. The general form of this series is

$$
\begin{equation*}
w=\sum_{m=1}^{M} \chi_{m}\left(\bar{x}_{s}\right) q_{m} \tag{10.4.81}
\end{equation*}
$$

Clearly, the specific form of the $\chi_{m}$ functions depends on how the structure is modeled.

The surface displacement enters the description of the pressure loading through the concept of virtual work. This quantity is unlike the actual work done by forces. Rather, it is the work that would result if each generalized coordinate were incremented by an arbitrary differential amount. These increments are denoted as $\delta q_{m}$ to distinguish

[^45]them from actual increments $d q_{m}$ that occur in the course of a response. Increments of the generalized coordinates lead to a virtual displacement. At the surface, this displacement is $\delta w$. The functions $\chi_{m}$ are set by the manner in which the structure has been modeled and therefore are unchanged. Thus, the virtual displacement is given by
\[

$$
\begin{equation*}
\delta w=\sum_{m=1}^{M} \chi_{m}\left(\bar{x}_{s}\right) \delta q_{m} \tag{10.4.82}
\end{equation*}
$$

\]

The resultant force exerted by the pressure acting on a patch of the surface is $p\left(x_{s}, t\right) d \mathcal{S}$ in the direction of $-\bar{n}\left(\bar{x}_{s}\right)$ and $\delta w$ at that location is positive if it is in the direction of $+\bar{n}\left(\bar{x}_{s}\right)$. Therefore, virtual work done by the pressure acting on $\mathcal{S}_{e}$ is

$$
\begin{equation*}
\delta W_{\text {pressure }}=-\iint_{\mathcal{S}_{e}} p\left(\bar{x}_{s}, t\right) \delta w d S \tag{10.4.83}
\end{equation*}
$$

Substitution of the cavity mode series for $p$ and the preceding expression for $\delta w$ into the virtual work gives

$$
\begin{equation*}
\delta W_{\text {pressure }}=-\sum_{n=0}^{N} \sum_{m=1}^{M} \iint_{\mathcal{S}_{e}}\left[p_{n} \Psi_{R, n}\left(\bar{x}_{s}\right)\right]\left[\chi_{m} \delta q_{m}\right] d S \tag{10.4.84}
\end{equation*}
$$

On the other hand, the fundamental definition of the elements of $\left\{F_{\text {pressure }}\right\}$ is that they are the coefficients in a canonical description of $\delta W_{\text {pressure }}$, specifically

$$
\begin{equation*}
\delta W_{\text {pressure }}=\sum_{m=1}^{M}\left(F_{\text {pressure }}\right)_{m} \delta q_{m} \tag{10.4.85}
\end{equation*}
$$

These alternative descriptions of $\delta W_{\text {pressure }}$ must be the same for any set of $\delta q_{j}$ values. This can only be true if corresponding coefficients of $\delta q_{m}$ are the same, which leads to

$$
\begin{gather*}
\left(F_{\text {pressure }}\right)_{m}=-\mathcal{S}_{e} \sum_{n=0}^{N} p_{n} \Lambda_{m, n}  \tag{10.4.86}\\
\Lambda_{m, n}=\frac{1}{\mathcal{S}_{e}} \iint_{\mathcal{S}_{e}} \chi_{m}\left(\bar{x}_{s}\right) \Psi_{R, n}\left(\bar{x}_{s}\right) d \mathcal{S}
\end{gather*}
$$

It follows that the pressure force vector is

$$
\begin{equation*}
\left\{F_{\text {pressure }}\right\}=-\mathcal{S}_{e}[\Lambda]\{p\} \tag{10.4.87}
\end{equation*}
$$

In turn, the structure's equations of motion become

$$
\begin{equation*}
[M]\{\ddot{q}\}+[K]\{q\}+\mathcal{S}_{e}[\Lambda]\{p\}=\{F\} \tag{10.4.88}
\end{equation*}
$$

The series description of surface displacement, Eq.(10.4.81), is also used to describe the surface acceleration term in the pressure equations. Substitution of the displacement series into Eq. (10.4.79) shows that

$$
\begin{equation*}
\iint_{\mathcal{S}_{e}} \Psi_{R, n} \ddot{w} d \mathcal{S}=\sum_{m=1}^{M}\left[\iint_{\mathcal{S}_{e}} \Psi_{R, n}\left(\bar{x}_{s}\right) \chi_{m}\left(\bar{x}_{s}\right) d \mathcal{S}\right] \ddot{q}_{m}=\mathcal{S}_{e} \sum_{m=1}^{M} \Lambda_{m, n} \ddot{q}_{m} \tag{10.4.89}
\end{equation*}
$$

When the preceding expression for the surface displacement is substituted into Eq. (10.4.79), the result is

$$
\begin{equation*}
\mathcal{V}\left[\ddot{p}_{n}+\left(\omega_{R, n}\right)^{2} p_{n}\right]-\rho_{0} c^{2} \mathcal{S}_{e} \sum_{m=1}^{M} \Lambda_{m, n} \ddot{q}_{m}=\Gamma_{n}(t), \quad n=0,1, \ldots, N \tag{10.4.90}
\end{equation*}
$$

where the source coefficients are

$$
\begin{equation*}
\Gamma_{n}=\rho_{0} c^{2} \iint_{\mathcal{S}_{v}} \Psi_{R, n}\left(\bar{x}_{s}\right) \frac{\partial}{\partial t} \hat{v}\left(\bar{x}_{s}, t\right) d \mathcal{S}+\subset_{0} \sum_{m} \Psi_{R, n}\left(\bar{x}_{m}\right) \dot{Q}_{m}(t) \tag{10.4.91}
\end{equation*}
$$

In keeping with the matrix representation of the structural equations of motion, the equations for the acoustical response may be written in that form. Stacking the pressure equations above the structural equations leads to a set of ordinary differential equations whose form is

$$
[\mathcal{M}] \frac{d^{2}}{d t^{2}}\left\{\begin{array}{l}
\{p\}  \tag{10.4.92}\\
\{q\}
\end{array}\right\}+[\mathcal{K}]\left\{\begin{array}{c}
\{p\} \\
\{q\}
\end{array}\right\}=\{\mathcal{F}\}
$$

The system matrices are

$$
\begin{gather*}
{[\mathcal{M}]=\left[\begin{array}{cc}
\mathcal{V}[I]-\rho_{0} c^{2} \mathcal{S}_{e}[\Lambda]^{\mathrm{T}} \\
{[0]} & {[M]}
\end{array}\right]} \\
{[\mathcal{K}]=\left[\begin{array}{c}
\mathcal{V}\left[\left(\omega_{R}\right)^{2}\right][0] \\
\mathcal{S}_{e}[\Lambda]
\end{array}\right]}  \tag{10.4.93}\\
\{\mathcal{F}\}=\left\{\begin{array}{l}
\{\Gamma\} \\
\{F\}
\end{array}\right\}
\end{gather*}
$$

where $\left[\left(\omega_{R}\right)^{2}\right]$ is a diagonal matrix whose elements are the rigid-cavity natural frequencies and [I] is an identity matrix. Thus, we have derived a set of coupled equations whose number equals the number of unknown pressure coefficients, $N+1$, plus the number of unknown displacement coefficients, $M$. The column vector $\{\Gamma\}$
represents source excitations stemming from injecting mass into the fluid, as well as from imposing a movement of the boundary. The displacement and pressure variables are coupled by the matrix [ $\Lambda$ ].

These equations cover most situations, but the case where some region in the wall is pressure-release is not addressed. Just as rigid mode series, Eq. (10.4.71), cannot give a nonzero normal velocity on the boundary, it cannot give a zero pressure on the boundary because the rigid-cavity modes are not zero there. An approximation of that condition may be obtained by replacing the pressure-release region with an extremely compliant structure, such as an extremely thin membrane whose surface tension is very low. However, doing so might cause the final equations to be ill-conditioned if the thickness and surface tension are too small.

How to solve the equations derived from Dowell's approximation depends on the type of response we seek. If the excitation is time harmonic, we may obtain the steadystate response from a complex variable representation. We set $p_{n}=\operatorname{Re}\left(P_{n} e^{i \omega t}\right)$, $q_{n}=\operatorname{Re}\left(\xi_{n} e^{i \omega t}\right), \Gamma_{n}=\operatorname{Re}\left(\hat{\Gamma}_{n} e^{i \omega t}\right)$, and $F_{n}=\operatorname{Re}\left(\hat{F}_{n} e^{i \omega t}\right)$, which leads to

$$
\left[[\mathcal{K}]-\omega^{2}[\mathcal{M}]\right]\left\{\begin{array}{c}
\{P\}  \tag{10.4.94}\\
\{\xi\}
\end{array}\right\}=\left\{\begin{array}{c}
\{\hat{\Gamma}\} \\
\{\hat{F}\}
\end{array}\right\}
$$

Resonances occur at frequencies for which the determinant $\left|[\mathcal{K}]-\omega^{2}[\mathcal{M}]\right|$ vanishes.
Solution of the differential equations in the time domain is a little more difficult. Coupled linear ordinary differential equations like Eq.(10.4.92) arise in mechanical and structural vibratory systems. In that area, the equations are commonly solved by modal analysis, in which the modes constitute the eigensolutions of the homogeneous equations. That formulation is not directly applicable here because the coefficient matrices are not symmetric. Consequently, a modal analysis would require consideration of left and right eigenvectors, which is the way in which the equations of motion for rotordynamic systems may be solved. ${ }^{15}$ The formulation of such solutions exceeds the scope of the present work.

The differential equations also may be solved by implementing a numerical algorithm, such as one of the Runge-Kutta versions. Most algorithms require that Eq.(10.4.92) be placed in first-order form. This is done by defining a state vector $\{Z\}$ that consists of all unknown variables and their first derivatives,

$$
\{Z\}=\left[\begin{array}{llll}
\{p\}^{\mathrm{T}} & \{q\}^{\mathrm{T}} & \{\dot{p}\}^{\mathrm{T}} & \{\dot{q}\}^{\mathrm{T}} \tag{10.4.95}
\end{array}\right]^{\mathrm{T}}
$$

The variables and their first derivatives may be extracted from $\{Z\}$ according to

$$
\left\{\begin{array}{l}
\{p\}  \tag{10.4.96}\\
\{q\}
\end{array}\right\}=[[I] \quad[0]]\{Z\}, \quad \frac{d}{d t}\left\{\begin{array}{l}
\{p\} \\
\{q\}
\end{array}\right\}=[[0] \quad[I]]\{Z\}
$$

[^46]where the size of the partitions equals the number of unknowns contained in $\{p\}$ and $\{w\}$, which is $N+M+1$. The derivative identity describes the fact that the derivative of the first expression equals the second, from which it follows that
\[

$$
\begin{equation*}
[[I] \quad[0]] \frac{d}{d t}\{Z\}=[[0] \quad[I]]\{Z\} \tag{10.4.97}
\end{equation*}
$$

\]

We also use Eq. (10.4.96) to describe the basic differential equations, Eq. (10.4.92), as

$$
\begin{equation*}
[[0] \quad[\mathcal{M}]]^{\mathrm{T}} \frac{d}{d t}\{Z\}+[[\mathcal{K}] \quad[0]]\{Z\}=\{F\} \tag{10.4.98}
\end{equation*}
$$

Stacking the derivative identity above this form yields

$$
\left[\begin{array}{l}
{[I][0]}  \tag{10.4.99}\\
{[0][\mathcal{M}]}
\end{array}\right] \frac{d}{d t}\{Z\}=\left[\begin{array}{cc}
{[0]} & {[I]} \\
-[\mathcal{K}] & -[0]
\end{array}\right]\{Z\}+\left\{\begin{array}{l}
\{0\} \\
\{F\}
\end{array}\right\}
$$

To solve these equations with mathematical software, one merely needs to provide the equation solver with the coefficient matrix multiplying $d\{z\} / d t$ and a function that evaluates the right side of the equality sign at an arbitrary instant based on the current value of $\{z\}$. Initial values required to begin the integration process are the values of $\{p\},\{w\},\{\dot{p}\}$, and $\{\dot{w}\}$ at $t=0$.

We have seen that Dowell's approximation leads to a set of differential equations that are readily solved and are quite comprehensive in the variety of systems to which the may be applied. However, they do have two flaws. The structural and fluid velocities are not inherently continuous in the normal direction at the interface where the media meet. This was observed at the outset, but another fault was encountered when the differential equations were assembled. The coupled equations are not symmetric. It was shown in Chap. 6 that the laws of acoustics are reciprocal. That is suppose a source is situated at one point and the response is measured at another. The same response would be measured at the original source location if the source were applied at the measurement point. It can be proven that this property requires that the matrix equations for the response be symmetric. This is not the nature of Eq. (10.4.92). Another consequence of the asymmetry of $[\mathcal{M}]$ and $[\mathcal{K}]$ arises in the computation of natural frequencies. These correspond to nonzero responses in the absence of excitation, so they are solutions of an eigenvalue problem,

$$
\left[[\mathcal{K}]-\omega^{2}[\mathcal{M}]\right]\left[\begin{array}{ll}
\{p\}^{\mathrm{T}} & \{q\}^{\mathrm{T}} \tag{10.4.100}
\end{array}\right]^{\mathrm{T}}=0
$$

It is possible that some eigenvalues $\omega^{2}$ will be negative, or they might occur as pairs of complex conjugate values, even though both matrices are real. Neither condition corresponds to a periodic motion. Such a prediction is erroneous, because periodic free vibration is the only possibility in any system that does not dissipate energy, as we saw in the development of the Rayleigh ratio.

Despite these flaws, results obtained from Dowell's method have been found to be quite usable, especially when the fluid is air. Indeed, the formulation has been
widely employed to predict cabin noise for aircraft and automobiles. The reason for this success is that the air presents a low impedance to movement of a solid structure. Consequently, the fluid's particle velocity at the interface will be much lower than it is somewhere in the interior of the cavity. (Such an assertion is not likely to be true for a heavy fluid like water.) In addition, issues regarding reciprocity seldom are at the forefront of the questions that the model equations must answer. In any event, the availability of a rather accessible simulation technique makes Dowell's approximation a very useful predictive tool for design and diagnostic tasks.

EXAMPLE 10.13 The sketch depicts a rectangular cavity in which both walls parallel to the plane of the sketch are rigid. At the top, a portion of the wall consists of a plate whose mass is $M_{1}$. The wall at the left side is another plate whose mass is $M_{2}$. Both plates are extremely stiff, so their movement is essentially a rigid body translation, vertically for the upper plate and horizontally for the side plate. The width of both plates perpendicular to the plane of the diagram is the same as the width of the waveguide, which is 300 mm . The upper plate is supported by spring $K$ and dashpot $D$, so this subsystem constitutes a one-degree-of-freedom oscillator. A horizontal force applied to the plate on the left is adjusted such that at any frequency $\omega$, it induces a translational velocity of $V_{0} \sin (\omega t)$, where $V_{0}=50 \mathrm{~mm} / \mathrm{s}$. All other regions of the wall are rigid. The value of $K / M_{1}$ has been selected such that the natural frequency of the oscillator in the absence of fluid loading would equal the lowest nonzero natural frequency of the cavity if all walls were rigid. The value of $D$ is set to give a 0.001 ratio of critical damping for the isolated oscillator. Both $M_{1}$ and $M_{2}$ are fifty times greater than the mass of the fluid within the cavity. Determine the pressure at three points: $x=90, y=37.5 \mathrm{~mm} ; x=180, y=75 \mathrm{~mm}$; and $x=360 / 2^{1 / 2}, y=150 / 2^{1 / 2} \mathrm{~mm}$, as a function of frequency up to $25 \%$ greater than the fourth natural frequency of the system. Also, determine the amplitude of the plate displacement within this range.


Figure 1.

## Significance

The analysis will demonstrate all of the operations required to implement Dowell's approximation without requiring that we formulate a complicated model of the
structure. Features of special interest are the way in which the discontinuous properties of the surface are handle and the influence of the structure on the overall response characteristics.

## Solution

The coordinate system we shall use is shown in the sketch. Each plate is stated to only translate in the direction normal to its surface. "Translation" means that all locations execute the same displacement in the $z$ direction. The motion of the plate at $x=0$ is the excitation, and a pressure field that is independent of $z$ may be sustained because the sidewalls are rigid. This observation permits formulation of a two-dimensional model, in which the pressure field depends only on the $x$ and $y$ coordinates. Correspondingly, integrals over a surface will feature a width factor, $W=0.3 \mathrm{~m}$.

The interface of the horizontal plate and the fluid, which constitutes $\mathcal{S}_{e}$, occupies $0.24<x<0.36, y=0.15 \mathrm{~m}$. The total area of this portion of the surface is $\mathcal{S}_{e}=$ $W(0.12)=0.036 \mathrm{~m}^{2}$. All points on this surface displace by the same amount $w(t)$, positive for downward movement (into the fluid). This corresponds to Eq. (10.4.81) with a single generalized coordinate and a surface displacement function that is one, that is,

$$
\begin{equation*}
w=\chi_{1} q, \quad \chi_{1}=1 \text { for } 0.24<x<0.36 \tag{10.4.101}
\end{equation*}
$$

The plate that is situated at $x=0$ undergoes a specified motion, and the remainder of the boundary is rigid. This means that all portions of the boundary excluding the oscillator's face constitute $\mathcal{S}_{v}$, with $\hat{v}=\operatorname{Re}\left(V_{0} e^{i \omega t}\right)$ on $x=0$ and $\hat{v}=0$ elsewhere. Consequently, integrals containing $\hat{v}$ will have a nonzero contribution from $x=0$ only.

A rigid-cavity modal series was formulated in Example 10.8. A variant of the single index sequencing scheme used there is somewhat simpler to formulate. Index $m$ is associated with cavity mode $(j, n)$, with $n$ being the number for $x$ dependence. The terms associated with each cavity mode may be defined in a double loop over $j=0, \ldots, J$ and $n=0, \ldots, N$. The associated value of $m$ is

$$
\begin{equation*}
m=n+1+j(N+1), \quad j=0,1, \ldots, J, \quad n=0,1, \ldots, N \tag{1}
\end{equation*}
$$

where the highest mode numbers in each direction are $J$ and $N$. Correspondingly, the rigid-cavity modal properties are

$$
\begin{gather*}
\Psi_{R, m}=B_{m} \cos \left(\frac{j \pi x}{L}\right) \cos \left(\frac{n \pi y}{H}\right), \quad m=1,2, \ldots,(J+1)(N+1) \\
\omega_{R, m}=c\left[\left(\frac{j \pi}{L}\right)^{2}+\left(\frac{n \pi}{H}\right)^{2}\right]^{1 / 2} \tag{2}
\end{gather*}
$$

where $H=0.15 \mathrm{~m}$ and $L=0.36 \mathrm{~m}$. The modes are normalized according to Eq. (10.3.8), which gives

$$
\begin{equation*}
B_{m}=\frac{2}{\left(1+\delta_{0, j}\right)^{1 / 2}\left(1+\delta_{0, n}\right)^{1 / 2}} \tag{3}
\end{equation*}
$$

There are no sources within the fluid, so the source coefficients only contain a contribution from the imposed motion of the plate on the left. Therefore, the surface integral in Eq. (10.4.91) for these coefficients extends only over that plate,

$$
\begin{align*}
\Gamma_{m, 1} & =\rho_{0} c^{2} \iint_{\mathcal{S}_{v}} \Psi_{R, m}\left(\bar{x}_{s}\right) \frac{\partial}{\partial t} \hat{v}\left(\bar{x}_{s}, t\right) d \mathcal{S} \\
& =\left.\rho_{0} c^{2} \int_{0}^{H} \Psi_{R, m}\left(\bar{x}_{s}\right)\right|_{x=0}\left(\omega V_{c} \cos (\omega t)\right)(W d y) \\
& =\rho_{0} c^{2} W \int_{0}^{H} B_{m} \cos \left(\frac{n \pi y}{H}\right)\left(\omega V_{0} \cos (\omega t)\right) d y=\rho_{0} c^{2} W H B_{m} \omega V_{0} \cos (\omega t) \delta_{n, 0} \tag{4}
\end{align*}
$$

This tells us that only the modes that are constant in the $y$ direction are directly excited.

The matrix [ $\Lambda$ ] mathematically couples the modes and the displacement of the upper plate. The coefficients obtained from Eq. (10.4.86) are

$$
\begin{align*}
\Lambda_{1, m} & =\left.\frac{1}{\mathcal{S}_{e}} \int_{2 L / 3}^{L}\left[\chi_{1}\left(\bar{x}_{s}\right) \Psi_{R, m}\left(\bar{x}_{s}\right)\right]\right|_{y=H}(W d x) \\
& =\frac{W}{\mathcal{S}_{e}} \int_{2 L / 3}^{L} B_{m} \cos \left(\frac{j \pi x}{L}\right) \cos (n \pi) d x  \tag{5}\\
& =\left\{\begin{array}{l}
\frac{W}{\mathcal{S}_{e}}\left(\frac{L}{j \pi}\right) B_{m} \sin \left(\frac{2 j \pi}{3}\right)(-1)^{n+1} \text { if } j \neq 0 \\
\frac{W L}{3 \mathcal{S}_{e}} B_{\ell}(-1)^{n} \text { if } j=0
\end{array}\right.
\end{align*}
$$

Note that each element of $[\Gamma]$ and $[\Lambda]$ is computed within the loops over $j$ and $n$.
The equations for the acoustical field are supplemented by those for the structure. In the present case, a single differential equation of motion governs the displacement variable $q$ of the oscillator. The only external force acting on this system is the resultant force exerted by the pressure field, which is obtained by integrating the pressure over the face of the piston. Because a positive $q$ has been defined to be positive for downward displacement, a positive pressure represents a negative resultant force. Thus, this force is

$$
\begin{align*}
F_{\text {pressure }} & =-\iint_{\mathcal{S}_{e}} p\left(\bar{x}_{s}, t\right) d \mathcal{S}=\left.\sum_{m=1}^{(J+1)(N+1)} \int_{2 L / 3}^{L} \Psi_{R, m}\left(\bar{x}_{s}\right)\right|_{y=H} p_{m}(W d y),  \tag{6}\\
& =-\mathcal{S}_{e} \sum_{m=1}^{(J+1)(N+1)} \Lambda_{1, m} p_{m}
\end{align*}
$$

The equation of motion for the oscillator is $M_{1} \ddot{q}+C \dot{q}+K q=F_{\text {pressure }}$. Substitution of Eq. (6) gives

$$
\begin{equation*}
M_{1} \ddot{q}+C \dot{q}+K q+\mathcal{S}_{e} \sum_{m=1}^{(J+1)(N+1)} \Lambda_{1, m} p_{m}=0 \tag{7}
\end{equation*}
$$

The mass is specified to be $M_{1}=50 \rho_{0} \mathcal{V}$, and $K$ is set by the requirement that the isolated natural frequency $\left(K / M_{1}\right)^{1 / 2}$ equals the fundamental nonzero natural frequency of the rigid-cavity. This frequency belongs to either the $(1,0)$ or $(0,1)$ mode. Because $L>H$, the former is the choice, so we set

$$
\frac{K}{M_{1}}=c^{2}\left(\frac{\pi}{L}\right)^{2}
$$

The fundamental nonzero frequency of the cavity is 472 Hz . The other parameter is $D$, which sets the ratio of critical damping to 0.001 . This ratio is such that division of Eq. (7) by $M_{1}$ converts the coefficient of $\dot{q}$ to $D / M_{1}=2$ (ratio) $\left(K / M_{1}\right)^{1 / 2}$, so $D=2(0.001)\left(K M_{1}\right)^{1 / 2}$. The values that result are

$$
M_{1}=M_{2}=0.972 \mathrm{~kg}, K=8.57\left(10^{6}\right) \mathrm{N} / \mathrm{m}, D=57.7 \mathrm{~N}-\mathrm{s} / \mathrm{m}
$$

The general frequency-domain equations are described in Eq. (10.4.94). To adapt them to the present analysis, we populate $\{P\}$ with the $P_{m}$ coefficients and replace the structural displacement vector $\{W\}$ with a single coefficient $\hat{q}$, which is the complex amplitude of $q$. Correspondingly, $[\Lambda]$ is a row vector formed from the coefficients $\Lambda_{1, \ell},[K]$ is the spring stiffness $K$, and $[M]$ is the piston mass $M_{2}$. The structural damping term gives rise to a term $i \omega C$ that adds to $K-\omega^{2} M$ in the oscillator's equation of motion. The frequency-domain equations that result are

The computational procedure begins by computing all quantities that do not depend on the frequency. This is done with a pair of nested loops in which $j$ ranges from 0 to $J$, within which $n$ ranges from 0 to $N$. Inside these loops, the vector index is $m=n+1+j(N+1)$. The $m$ th element of $\left[\left(\omega_{R}\right)^{2}\right],[\Lambda]$, and $\{\tilde{Q}\}$ are evaluated. Then, a loop over frequency formulates and solves Eq. (8) over a range of frequencies. The column vector $\{P\}$ of complex pressure amplitudes at each $\omega$ is saved as a column of an rectangular matrix [ $P_{\text {all }}$ ], and the corresponding complex displacement amplitude $\zeta$ is saved as an element of a vector.

One response that is requested is the pressure at the midpoint of the cavity. Equation (10.4.71) indicates that at any point $\left(x_{f}, y_{f}\right)$ the pressure is

$$
\begin{equation*}
P=\sum_{m=1}^{(J+1)(N+1)} P_{m} B_{m} \cos \left(\frac{j \pi x_{f}}{L}\right) \cos \left(\frac{n \pi y_{f}}{H}\right) \tag{9}
\end{equation*}
$$

The upper frequency limit is specified to be $25 \%$ greater than the system's fourth nonzero natural frequency. We do not know what that frequency is, so let us assume that the natural frequencies are close to those of the rigid cavity. The lower natural frequencies correspond to $n=0,1$, or 2 and $j=0,1$, or 2 . The values are tabulated below.

$$
\begin{array}{l||ccc} 
& n=0 & n=1 & n=2 \\
\hline j=0 & 0 & 7121 & 14242 \\
j=1 & 2967 & 7714 & 14548 \\
j=2 & 5934 & 9269 & 15429
\end{array}
$$

The fourth nonzero value belongs to the $(1,1)$ mode. Setting the cutoff to $25 \%$ above that value gives $9643 \mathrm{rad} / \mathrm{s}$. We will round it up to 1.6 kHz .

We must decide how many modes to include. One guideline comes from making sure that the highest natural frequency of the model is well above the maximum excitation frequency. This will not be a difficult criterion to meet. A more stringent requirement is that the number of modes should be sufficient to recreate in a modal series the impedance discontinuity of the upper wall, but that provides no definitive guideline. It is reasonable to set the highest mode numbers $J$ and $N$ such that the $(0, N)$ and $(J, 0)$ modes have comparable natural frequencies and that these frequencies are at least twice the highest excitation frequency. We shall try $J=10$, which leads to $N=\operatorname{ceil}(J(H / L))=5$. Because of the vagueness of these guidelines, the computed results were verified by reducing $J$ to 5 and $N$ to 3 . Doing so resulted in barely perceptible alterations of the graphed data.

The upper graph in Fig. 2 describes the pressure responses at the prescribed locations. The lower graph shows the corresponding plate displacement. The vertical lines mark the rigid-cavity natural frequencies. The peaks that occur are very close to these values. This attribute is a consequence of the oscillator having a large mass. It appears to the fluid to be a nearly rigid obstacle. Consequently, the resonances in the frequency responses occur close to the natural frequencies of the rigid-cavity modes. However, the pressure at the midpoint is very small when $\omega$ is close to the natural frequency of the $(1,0)$ rigid body mode. The pressure at the other locations seems to be nullified at that frequency. The explanation for this feature lies in the concept of a tuned vibration absorber, in which a low-mass oscillator is mounted on a system. The natural frequency of the attachment is set equal to that of the unadorned system. The result is that the previously unadorned part of the system does not respond at that frequency. The piston-spring oscillator on the upper wall serves the same purpose, except that its mass is much greater than the fluid.


Figure 2.

It still remains to explain why some locations do not show evidence of resonances that are seen at other locations. This behavior is a corollary of the proximity of resonances to the rigid-walled behavior. Consequently, at each resonant frequency $\omega_{m} \approx\left(\omega_{R}\right)_{m}$, the pressure coefficient $P_{m}$ for the resonant mode is much larger than the other coefficients. When the rigid-cavity modal series is evaluated at this frequency, the $m$ th mode will make the dominant contribution. However, if the corresponding mode function $\Psi_{R, m}$ is small at a specific location, then the pressure there will be much smaller than it is elsewhere. In the case of the midpoint, where $x=L / 2$ and $y=$ $H / 2$, we have $\Psi_{R, m}=B_{\ell} \cos (j \pi / 2) \cos (n \pi / 2)$. Thus, if a resonance corresponds to a cavity mode for which $j$ or $n$ is odd, then that resonance will not be observed at the midpoint. Because $x=360 / 2^{1 / 2}$ and $y=150 / 2^{1 / 2}$ are irrational numbers, none of the rigid-cavity mode functions are zero at that point. Consequently, every natural frequency is manifested in a peak of the pressure at that point.

Another interesting feature is the narrow peak and adjacent null that occurs at some frequencies. These all correspond to modes for which $n=1$. The explanation lies in the source coefficients $\hat{Q}_{\ell}$, which represent the excitation. According to Eq. (4), these values are nonzero only if $n=0$. Thus, only the modes that correspond to $n=0$ are excited directly. If the resonance is of a cavity mode for which $n \neq 0$, the excitation is much weaker than it is for $n=0$ modes. The drastic change from a maximum to near-null results from the fact that crossing a resonance results in $180^{\circ}$ phase shift for the $P_{\ell}$ coefficients that are resonant, so instead of adding to the contribution of the nonresonant coefficients, they subtract.

The last effect of note is the fact that the oscillator displacement shows no evidence of the $(3,0)$ resonance, even though the pressure does. This feature is a consequence of the fact that the length of the piston in the $x$ direction is $L / 3$. Equation (5) indicates that the $\Lambda_{1, \ell}$ values are zero if $\ell$ corresponds to a mode for which $j$ is an integer multiple of 3 . This is the mathematical consequence of the fact that $\Lambda_{1, \ell}$ represents
the interaction between the piston and a mode and that interaction averages out over the length of the piston if the pressure along it varies as $\cos (3 \pi x / L), \cos (6 \pi x / L)$, etc

Many of the features discussed in the preceding arise because the structural mass $M_{2}$ is large. A system in said to be a case of "light fluid loading" if the structure's mass is large compared to the fluid's. Decreasing the mass of the upper plate leads to system with "heavy fluid loading," although it would be more accurate to say that it is a system with low structural impedance. Dowell's approximation might not be applicable in this parameter range. Nevertheless, carrying out the computation with the plate mass decreased by a factor of 50 , so that it equals the mass of the fluid within the cavity and enables us to see how the mass ratio affects the response. Although the plate mass is decreased, the value of $K / M_{2}$ is maintained. This has the effect of increasing the importance of the term containing [ $\Lambda$ ] in the lower part of Eq. (8). As a result, coupling of the $P_{m}$ coefficients is stronger. Figure 3 shows the pressure responses for this case. The oscillator now is a less substantial obstacle to movement of the fluid, so the cavity's behavior differs considerably from a cavity with rigid walls. One consequence is that the resonant frequencies are noticeably different from the natural frequencies of the rigid-cavity. A by-product is that resonances are more likely to be seen at all points. Stronger coefficient coupling also accounts for the increased occurrence of displacement nulls at various frequencies.


Figure 3.

An overview indicates that Dowell's approximation is readily implemented to describe fairly complicated systems. The most difficult operation in such cases is likely to be determination of the rigid-cavity modes, which are a prerequisite to the formulation. That was not aan issue for the box cavity, nor would it be for any of the rectangular, cylindrical, or spherical cavities we have studied.

### 10.5 Homework Exercises

Exercise 10.1 An exponential horn whose length is $L$ is closed at its narrow end and open to the atmosphere at its wide end. The cross section is sufficiently small that only plane waves can propagate within it. Derive the characteristic equation whose roots are the horn's natural frequencies. Then, evaluate these frequencies for a horn whose length is $L=\pi / k$ with $b=0.08 / L$. Compare these values to the natural frequencies for plane waves in a uniform waveguide whose length is $L$ with one end open and the other closed.

Exercise 10.2 The sketch depicts a two-dimensional model of a box filled with a liquid. The sides are rigid, and the top is open to the atmosphere. The bottom is bonded to the ground, so any vertical motion of the ground is replicated by the bottom. In the situation of interest, vibration of a nearby machine induces a vertical displacement $w=W \cos (2 \pi x / L) \sin (\omega t)$. Derive an expression for the pressure at any location within the tank. Use that expression to evaluate the vertical velocity $v_{y} /(W \omega)$ at the midpoint of the free surface for frequencies in the range $0<k L \leq 4 \pi$. The aspect ratio is $H / L=0.6$.


Exercises 10.2, 10.3, and 10.4.

Exercise 10.3 Consider the situation in which a polymer sheet is placed on the top of the liquid of the box in Exercise 10.2. The result is that the specific impedance at this surface $\zeta=0.8-3 i$. Evaluate the vertical velocity $\left|v_{z}\right| /(W \omega)$ at the midpoint of the free surface for frequencies in the range $0<k L \leq 4 \pi$. Then, compare this response to that which would be obtained with a pressure-release condition, $\zeta=0$ The aspect ratio is $H / L=0.6$.

Exercise 10.4 The system of interest is the one in Exercise 10.2, with the sole difference that the vertical wall at $x=L$ is an intertance, with $\zeta=0.25 i$. The aspect ratio is $H / L=0.6$. The excitation frequency is the lowest value at which the system would resonate if $x=L$ were rigid. Evaluate the vertical velocity $\left|v_{z}\right| /(W \omega)$ at the midpoint of the free surface.

Exercise 10.5 A container whose sides are otherwise rigid is excited by an oscillation in which the vertical wall at $x=0$ pivots about its horizontal centerline. The
result is that its normal velocity is $v_{x}=v_{0}(y / H-0.5) \cos (\omega t)$. The dimensions of the container are $L, H$, and $W$ in the $x, y$, and $z$ directions, respectively. Derive an expression that describes the pressure field within the container.

Exercise 10.6 A rectangular water tank has sides that are composed of very thick glass panels that may be taken to be rigid. The top, $y=H$, is open to the atmosphere, so it is pressure-release. The bottom, $y=0$, is a thin steel plate that flexes because of ground vibration. The consequence is that the plate undergoes a vibration normal to its surface, specifically, $v_{z}=v_{0} \sin (\pi x / L)$. The dimensions of the bottom plate are $L \times W$. Derive an expression for the pressure $p /\left(\rho_{0} c v_{0}\right)$ at an arbitrary location within the tank.

Exercise 10.7 The normal velocity at the left side of a water tank, $x=0$, is harmonic with a parabolic dependence on $y$ given by $v_{z}=v_{0}\left(y / b-y^{2} / b^{2}\right) \sin (\omega t)$ for $0<y<b$. The opposite side, $x=a$, is locally reacting, with local impedance $\zeta_{a}=0.6-0.3 i$. The sides $z=0$ and $z=d$ are rigid, as is the bottom, $y=0$, whereas $y=b$ is pressure-release. (a) Derive an expression for the pressure at any location. (b) The dimensions are $a=10 \mathrm{~mm}, b=5 \mathrm{~mm}$, and $d=3 \mathrm{~mm}$. Evaluate the pressure at the midpoint $x=a / 2, y=b / 2$, and $z=d / 2$. The frequency range is $k a \leq 16$. Are resonances evident in this frequency response?

Exercise 10.8 An engineered material has been installed in the ceiling of a closed room. It has the property of behaving like a continuous sheet of bricks, so that it is locally reacting, with an impedance of $0.1 k L$. The wall at $x=0$ undergoes a translational oscillation in which $v_{x}=v_{0} \sin (\omega t)$. The other walls and the floor are rigid. (a) Derive an expression for the pressure at an arbitrary location. (b) Identify the natural frequencies by examining the expression for the pressure. (c) The proportions of the room are $L: H: W=5: 3: 4$. Evaluate the pressure $|p| /\left(\rho_{0} c v_{0}\right)$ at the middle of the room as a function of frequency for $k a<12$.

Exercise 10.9 The walls, floor, and ceiling of a reverberant room are essentially rigid. Its height is $H$, and its horizontal dimensions are $L \times W$. It is excited by a loudspeaker that emits a constant amplitude at frequency $\omega$. The loudspeaker is situated at the middle of the ceiling. It is sufficiently small that it may be considered to be a point source having volume velocity is $\operatorname{Re}(\hat{Q} \exp (i \omega t))$. Derive an expression for the pressure at the middle of the room.

Exercise 10.10 The sound pressure at the middle of the reverberant room in Exercise 10.9 has been measured. At 1 kHz , it is $92 \mathrm{~dB} / / 20 \mu \mathrm{~Pa}$. Determine the volume velocity $\hat{Q}$. The room dimensions are 3 m high and $5 \mathrm{~m} \times 4 \mathrm{~m}$ horizontally.

Exercise 10.11 A small spherical object has been placed at the center of a spherical cavity filled with water. The outer wall of the cavity executes a radially symmetric vibration whose amplitude is fixed at $0.2 \mathrm{~m} / \mathrm{s}$, but the frequency is variable in the range from 1 to 10 kHz . The surface of the spherical object is locally reacting with $\zeta=0.6-0.4 i$. The radii are $a=800 \mathrm{~mm}, b=10 \mathrm{~mm}$. The region between the
surfaces is filled with water. Determine the frequency at which the pressure on the surface of the object is maximum. What are the pressure amplitudes on the inner and outer surfaces at this frequency?

Exercise 10.12 A transducer is flush mounted on the inner surface of a spherical cavity whose walls otherwise are rigid. The transducer is very small, so it may be approximated as a point source whose volume velocity is $\operatorname{Re}(\hat{Q} \exp (i \omega t))$. Derive an expression for the pressure at the center of the cavity.

Exercise 10.13 A cylindrical tank filled with oil has a piston transducer situated concentrically at $x=0$. The radius of the tank is $a$, and the radius of the piston is $a / 2$. Consequently, the resulting velocity at $x=0$ is $v_{x}=v_{0} \cos (\omega t)$ if $R<a / 2$, $v_{x}=0$ if $R>a / 2$. The other surfaces are rigid. Derive an expression for $p /\left(\rho_{0} c v_{0}\right)$ at an arbitrary location. Then, use that expression to obtain an expression for the acoustic force acting on the wall at $x=L$. Evaluate $\left|p /\left(\rho_{0} c v_{0}\right)\right|$ at $x=L, R=0$ and $|F| /\left(\pi a^{2}\right)$ as functions of $k a<15$. Parameters are $\rho_{0}=910 \mathrm{~kg} / \mathrm{m}^{3}, c=1460 \mathrm{~m} / \mathrm{s}$, $L=4 \mathrm{~m}$, and $a=0.25 \mathrm{~m}$.

Exercise 10.14 Exercise 9.27 described the axial velocity at the near end of a cylindrical waveguide when the termination is a misaligned disk that rotates about the centerline at angular speed $\Omega$. The velocity at a point whose cylindrical polar coordinates are $(R, \theta, x=0)$ is well described as $v_{x}=\Delta \Omega R \sin (\Omega t-\theta)$, where $\Delta$ is the angle between the normal to the disk and the centerline. Rather than taking the far end to be nonreflecting, as was done in Exercise 9.27, suppose that the end at $x=L$ is open to the atmosphere, which corresponds to a pressure-release termination. The cylindrical wall is rigid. Derive an expression for the pressure at an arbitrary location.

Exercise 10.15 The fluid contained in a cylinder whose length is $L$ and diameter is $2 a$ is excited by harmonic transverse vibration of the cylindrical wall. The velocity amplitude is axisymmetric with a variation in the axial direction as a single lobe of a sine, so that $v_{R}=v_{0} \sin (\pi x / L) \sin (\omega t)$. The ends of the cavity are rigid. Derive an expression for the pressure amplitude $|P| /\left(\rho_{0} c v_{0}\right)$ at an arbitrary location. Then, evaluate this expression at the midpoint, $x=L / 2, R=0$, as a function of frequency in the range $k a<12$. The aspect ratio is $L / a=5$.

Exercise 10.16 A cylindrical tank was filled with methyl alcohol ( $c=1103 \mathrm{~m} / \mathrm{s}$, $\rho_{0}=792 \mathrm{~kg} / \mathrm{m}^{3}$ ). A small transducer was mounted concentrically at one end. The far end was closed with mylar, which well fits a pressure-release model, whereas the cylindrical wall was lined with a special insulating coating. The radius of the tank was 3 m . The general analysis of the pressure field in a cavity indicates that in a nearly resonant situation the spatial distribution is essentially proportional to the resonant mode. With this in mind, the transducer was swept through a range of frequencies, with the result that 154.3 Hz was identified as the lowest frequency at which the interior field attains a peak amplitude. At this frequency, the pressure amplitude along a transverse line was measured. The data that was collected is described by the following table. The pressure waveforms at these locations were found to be
in-phase. (a) Determine the transverse wavenumber $\mu$ by fitting the measured pressure to the analytical expression for the pressure field. (b) Determine the dependence of the pressure on the axial distance measured from the driven end. (c) Determine the specific impedance of the liner material. (d) Determine the length of the cylinder.

$$
\begin{array}{c||r|r|r|r|}
R / a & 0.2 & 0.4 & 0.6 & 0.8 \\
\hline P(\mathrm{kPa}) & 310 & 268 & 203 & 125
\end{array}
$$

Exercise 10.17 The sketch depicts a rectangular enclosure in which a small source is situated at its midpoint. The source's volume velocity is $Q(t)=\operatorname{Re}(\hat{Q} \exp (i \omega t))$, where $\hat{Q}$ is independent of the frequency $\omega$. The enclosed fluid is oil $\left(\rho_{0}=910 \mathrm{~kg} / \mathrm{m}^{3}\right.$, $c=1460 \mathrm{~m} / \mathrm{s}$. All sides of the enclosure are rigid, except that the closure at the left end is a lightly tensioned membrane. The fluid on the other side of the membrane is air, so the membrane effectively constitutes a pressure-release termination. Furthermore, the left end is flush mounted with a very large rigid baffle, so the vibration of the membrane causes sound to be radiated into the air. In other words, this system acts like an unconventional type of loudspeaker. The primary property of a low-frequency loudspeaker is its volume velocity. Determine the magnification factor $\left|\hat{Q}_{\mathrm{mem}} / \hat{Q}\right|$ for the vibrating membrane's volume velocity. The frequency range of interest is zero to 2 kHz , and $L=0.5 \mathrm{~m}, W=0.25 \mathrm{~m}$.


Exercises 10.17

Exercise 10.18 Solve Exercise 10.17 for the case where the cavity is a cylinder. As described there, the closure at the left end is a pressure-release membrane and the other surfaces are rigid. The length is $L=0.5 \mathrm{~m}$, and the radius is $a=0.15 \mathrm{~m}$.

Exercise 10.19 The sketch shows two oppositely phased line sources arranged horizontally relative to the middle of the water tank. The sides and bottom are essentially rigid. Consider the situation where the line sources are harmonic, with volume velocities per unit length $\pm \operatorname{Re}(\hat{Q} \exp (i \omega t))$. Derive an expression for the pressure field when the horizontal separation $s$ is finite. (b) Take the limit as $s \rightarrow 0$ with $\hat{Q} s$ held constant at $D$ and thereby obtain the pressure field for a dipole line source at the center. (c) Evaluate the pressure distribution (amplitude and phase) along the bottom when $\omega=0.95 \omega_{1}$ and $\omega=\left(\omega_{1}+\omega_{2}\right) / 2$, where $\omega_{1}$ and $\omega_{2}$ are the two smallest
natural frequencies of modes that are excited by the dipole line source. The depth to length ratio is $H / L=0.4$.


Exercises 10.19

Exercise 10.20 A harmonic point source whose volume velocity is $\operatorname{Re}\left(Q_{0} \exp (i \omega t)\right)$ is situated at the midpoint of a cylindrical cavity. The ends at $x=0$ and $x=L$ are pressure-release, whereas the cylindrical wall, $R=a$, is rigid. Derive an expression for the pressure amplitude at an arbitrary point on the cylinder's axis. Evaluate the pressure amplitude $\left|P a^{2} /\left(\rho_{0} c Q_{0}\right)\right|$ as a function of distance along this axis in the case where $L=2 a$. Perform this evaluation for cases where the frequency is half the cavity's fundamental natural frequency and $10 \%$ greater than that frequency.

Exercise 10.21 The sketch shows a cylinder whose cross section is one-third of a full circle. All surfaces are rigid. Derive expressions for the natural frequencies and mode shapes of the fluid contained within this cavity.


Exercises 10.21.

Exercise 10.22 The wall of a spherical cavity is rigid. The volume of the fluid region is the same volume as the rectangular cavity in Example 10.7. Evaluate all natural frequencies of axisymmetric modes that are below 3 kHz . Graph the data by plotting the frequencies at a fixed polar harmonic against the root number.

Exercise 10.23 The sketch shows a bowl that is filled with water. The hemispherical wall is rigid, whereas the surface is free. Determine the axisymmetric mode functions of the water contained in the bowl.


Exercises 10.23

Exercise 10.24 The system of interest is a rigid-walled sphere whose radius is $a$. A point source whose volume velocity is $\operatorname{Re}(\hat{Q} \exp (i \omega t))$ is situated at the center. The pressure field may be described in terms of a series of the cavity's natural modes. An alternative analysis is suggested in Sect.6.4.2. According to it, the field may be considered to be the superposition of the field generated by the source in free space and a field whose properties are selected to satisfy the boundary conditions, which are violated by the free-space field. In the present case this condition is that the radial particle velocity should be zero at $r=a$. The task here is to formulate both solutions and then compare them by evaluating the dependence of the pressure on the radial distance from the center when $k a=12$. Part of the comparison should be an assessment of convergence of the modal series as the number of modes is increased.

Exercise 10.25 A harmonic point source whose volume velocity is $\operatorname{Re}\left(Q_{0} \exp (i \omega t)\right)$ is situated at radial distance $r=b$ from the center of a spherical cavity whose radius is $a$. In order to exploit axisymmetry properties, the radial line through the source is defined to be $\psi=\pi$. The wall of the cavity is rigid. Derive an expression for the pressure amplitude due the point source. Evaluate that expression to determine the radial dependence of the pressure amplitude $\left|P a^{2} /\left(\rho_{0} c \hat{Q}\right)\right|$ along the radial lines $\psi=0$ and $\psi=\pi / 2$. Parameters for these evaluations are $b=2 a / 3$ and $k a=5$.

Exercise 10.26 An exponential horn is closed at the narrow end and open to the atmosphere at the wide end. Use a series of one, two, and three Ritz basis functions to estimate the natural frequencies $k_{n} L$ in the case where the exponential growth factor is $\beta=0.7$. Compare the results to those obtained by solving the Webster horn equation. It may be assumed that the complex pressure amplitude depends only on the axial position $x$.

Exercise 10.27 The apex angle (side to side) of a cone is $10^{\circ}$. The cone is open at its a large end, and its wall is rigid. Use a series of one to four basis functions Ritz basis to the natural frequencies $k_{n} L$ to determine the natural frequencies of modes whose amplitude depends only on the radial distance from the apex.

Exercise 10.28 The apex angle (side to side) of a cone is $30^{\circ}$. The cone is open at its a large end, and its wall is rigid. The task here is to employ spherical coordinates $r, \psi, \theta$ centered at the apex to formulate the Rayleigh-Ritz equations. (a) Identify a set of basis functions that are suitable for identifying the natural frequencies of modes that are axisymmetric relative to the cone's axis, but vary in both the radial and polar direction. (b) Use a series of six of the Ritz basis in Part (a) to identify the nondimensional natural frequencies $k_{n} L$.

Exercise 10.29 The apex angle (side to side) of a cone is $30^{\circ}$. The cone is open at its large end, and its wall is rigid. The task here is to employ cylindrical coordinates $R, \theta, z$ to formulate the Rayleigh-Ritz equations. The $z$-axis should coincide with the cone's axis, and the origin should be at the cone's apex. (a) Identify a set of basis functions that are suitable for identifying the natural frequencies of modes that are axisymmetric relative to the cone's axis, but vary in both the radial and polar direction. (b) Use a series of six of the Ritz basis in Part (a) to identify the nondimensional natural frequencies $k_{n} L$.

Exercise 10.30 The cavity in the sketch is a sector of the region between concentric cylinders. It is filled with water. All walls are rigid. Use the Rayleigh-Ritz method to estimate the cavity's natural frequencies. Carry out the analysis using from one to eight basis functions.


Exercises 10.30

Exercise 10.31 Suppose that rather than being the interface with an elastic structure, $\mathcal{S}_{e}$ in the development of Dowell's approximation is the interface with a locally reacting material whose impedance is position dependent, $Z\left(\bar{x}_{s}\right)$. What are the corresponding differential equations governing the coefficients of the pressure series?

Exercise 10.32 A piston of mass $M$, which closes the waveguide in the sketch, is restrained by a spring whose stiffness is $K$. The termination at the other end
is rigid. A harmonically varying force $F$ is applied to the piston. Use Dowell's approximation to determine the pressure amplitude on the face of the piston as a function of the nondimensional frequency $k L<20$. Parameters for the evaluation are $M /\left(\rho_{0} \mathcal{A} L\right)=5$ and $(K / M)^{1 / 2}=2 \pi c / L$. Also determine the lowest five natural frequencies of the system.


Exercises 10.32

Exercise 10.33 The sketch depicts a closed tank filled with nitrogen, $\rho_{0}=$ $1.165 \mathrm{~kg} / \mathrm{m}^{3}, c=354 \mathrm{~m} / \mathrm{s}$. A plate that fully closes the right end is supported by a spring. The mass per unit surface area of the plate is $25 \mathrm{~kg} / \mathrm{m}^{2}$ and the spring stiffness is $500 \mathrm{MN} / \mathrm{m}$. The left end is driven by a plate that extends over half the height of the tank. Both plates translate as rigid bodies, and both extend horizontally across the full width of the waveguide. All other surfaces are rigid. Consequently, the system is two-dimensional in the plane of the sketch. The cavity is excited by harmonic vibration of the left plate at 1.6 kHz , such that $v=20 \sin (\omega t) \mathrm{mm} / \mathrm{s}$. Determine the pressure distribution along the bottom and the amplitude of the displacement of the plate at the right end.


Exercises 10.33.

Exercise 10.34 The region enclosed by the cylinder is filled with an ideal fluid. The left end is terminated by a concentric piston whose radius is $a / 2$ and whose mass is $M$. The remainder of the left end is rigid. Translational displacement $w$ of the piston is resisted by spring $K$. The cavity is excited by a piston at the right end whose radius is $a$. As a result, the axial velocity at this end is a specified function $v_{L}(t)$ for $0 \leq R \leq a$. (a) Use Dowell's approximation to derive a set of differential equations whose solution would be the pressure coefficients for the acoustic field and the displacement $w$ of the left piston. (b) The parameters are $M /\left(\rho_{0} \mathcal{A} L\right)=$

20, $L / a=3$, and $(K / M)^{1 / 2}=2 \pi c / L$. The imposed vibration at the right end is $v_{L}=v_{0} \cos (\omega t)$. Determine and graph the amplitude of the displacement $w$ and the pressure at the center of the right end as a function of frequency for $0<\omega<6 \pi c / L$. (c) For the parameters in Part (b), determine the first eight natural frequencies of the axisymmetric modes.


Exercises 10.34

## Chapter 11 <br> Geometrical Acoustics

Our explorations thus far have analyzed sound fields from a global perspective, in which field equations and boundary conditions for the entire system were satisfied concurrently. The concepts in this chapter take a localized view to the question of how a signal is modified as it propagates. It constructs rays, which are the paths along which a specific phase propagates through space, and wavefronts, which are surfaces of constant phase at a specific instant. Some aspects of the formulation are an application of differential geometry, so the method is referred to as geometrical acoustics. Because ray paths are the primary property that is derived, an alternative name for the formulation is ray acoustics. The concept is limited to high-frequency signals, but it is quite general otherwise.

A common application of geometrical acoustics is analysis of propagation through a heterogeneous fluid, in which the ambient density $\rho_{0}$ and pressure $p_{0}$ are functions of position. However, the ambient properties shall be considered to be independent of time, on the basis that they vary much more slowly than the time scale for the acoustic signal. Analysis of propagation through such a heterogeneous medium is a core task for environmental acoustic studies of the atmosphere and the ocean. The literature on geometrical acoustics and its application to propagation in the ocean and atmosphere is rich. The objective here is to introduce the fundamental concepts. The texts written by Brekhovskikh and Lysanov, ${ }^{1}$ Pierce, ${ }^{2}$ and Jensen et al. ${ }^{3}$ are recommended starting points for further studies.

In order to emphasize development of the basic concepts, we shall add another restriction to the systems we shall consider. Situations where the fluid is not quiescent

[^47]in its ambient state are excluded. Winds in the atmosphere might be a reasonably large fraction of the speed of sound. Thus, propagation of sound in the atmosphere is not covered here. In a strict sense, the ocean also is not at rest, but the speed of ocean currents is much less than the speed of sound. Although flow effects are omitted from the present development, the principles and techniques that are derived provide the foundation for study of sound propagation in the atmosphere.

### 11.1 Basic Considerations: Wavefronts and Rays

We begin with the general representation of a harmonic signal that depends arbitrarily in position, which is $p(\bar{x}, t)=\operatorname{Re}\left(F(\bar{x}) e^{i \omega t}\right)$. The polar representation of the complex amplitude may be written as $F(\bar{x})=P(\bar{x}) \exp (-i \Delta(\bar{x}))$, where $P(\bar{x})$ and $\Delta(\bar{x})$ are real. The geometrical acoustics approximation defines the nature of $\Delta(\bar{x})$. It considers any signal to be locally planar, with $\bar{e}(\bar{x})$ being the propagation direction and $c(\bar{x})$ being the local sound speed. Such a signal corresponds to $\Delta(\bar{x})=\bar{x} \cdot \bar{e}(\bar{x}) / c(\bar{x})$. (No reference time lag is incorporated because we can define $t=0$, however, we wish to obtain a pure sine or pure cosine, or any other phase.) Correspondingly, the geometrical acoustics ansatz for the pressure field is

$$
\begin{equation*}
p(\bar{x}, t)=P(\bar{x}) \operatorname{Re}\left(e^{i \omega(t-\bar{x} \cdot \bar{e}(\bar{x}) / c(\bar{x}))}\right) \tag{11.1.1}
\end{equation*}
$$

The fact that the amplitude function $P(\bar{x})$ is real will be crucial to later developments.
To identify how the preceding description is related to wavefronts, we begin with the fundamental definition that a wavefront is a surface along which the signal at the same instant has the same phase. A surface that is fixed in space is the locus of points at which a function of the position coordinates has a constant value. A moving surface may be obtained by replacing the constant value by a function of time. Thus, an equation whose form is $f(\bar{x})=t$ defines a surface that moves through space.

The phase $\theta$ of the signal in Eq. (11.1.1) is the coefficient of $i$ in the complex exponential, so we have

$$
\begin{equation*}
\frac{\theta}{\omega}=t-\frac{\bar{x} \cdot \bar{e}(\bar{x})}{c(\bar{x})} \tag{11.1.2}
\end{equation*}
$$

If we fix $t$, then the locus of points on a wavefront of constant specified $\theta$ may be found by solving the preceding for all possible $\bar{x}$. Thus, this relation is the equation defining the wavefront of constant phase $\theta$. It is useful for further developments to collect the position-dependent terms in a single function $\tau(\bar{x})$, whose dimensionality is time,

$$
\begin{equation*}
\tau(\bar{x}) \equiv \frac{\bar{x} \cdot \bar{e}(\bar{x})}{c(\bar{x})} \tag{11.1.3}
\end{equation*}
$$

The description of a wavefront provided by Eq. (11.1.2) actually has two interpretations. The values of $\bar{x}$ on a wavefront, and therefore wavefront's shape and overall
position, depend of the value of $t-\theta / \omega$. If we hold the phase at a constant $\theta$ and vary $t$, we will see the wavefront flow through space as time evolves. Alternatively, we can fix $t$ and construct a family of wavefronts at that instant corresponding to different values of the $\theta$. In other words, wavefronts may be considered to be a set of surfaces at a specific instant or a single surface that moves through space. Both views are equivalent, because only the value of $t-\theta / \omega$ is important. This duality tells us that everything we need to know may be found by following the wavefront corresponding to zero phase. Therefore, we take the wavefront equation to be

$$
\begin{equation*}
\tau(\bar{x})=t \tag{11.1.4}
\end{equation*}
$$

After we have described the zero-phase wavefront at an instant $t$, the wavefront along which the phase is $\theta \neq 0$ will occupy the same location at the instant $t^{\prime}=t+\theta / \omega$. $^{4}$

The description on Eq.(11.1.1) is quite general. The geometrical acoustic approximation is based on a fundamental assumption regarding spatial dependence of the amplitude and phase. For a simple plane wave, the phase is $\theta=\omega \bar{x} \cdot \bar{e} / c=k x$ so that $\nabla \theta=k \bar{e}=(2 \pi / \lambda) \bar{e}$. In view of the definition of $\theta$ in Eq.(11.1.2), we may consider $\nabla \theta$ to be proportional to the reciprocal of a local wavelength. Thus, variation of the phase over a small spatial scale implies that $\nabla \theta$ is relatively large. With this observation in mind, let us differentiate Eq.(11.1.1) to determine the pressure gradient,

$$
\begin{equation*}
\nabla p \equiv[\nabla P-i P \nabla \theta] \operatorname{Re}\left(e^{i(\omega t-\theta(\bar{x}))}\right) \tag{11.1.5}
\end{equation*}
$$

An overview of the various signals we have considered indicates that in many cases, the wavelength scale is much smaller than the spatial scale over which the pressure amplitude varies. A geometric acoustics analysis is based on the a priori assumption of this behavior. More definitively, it is required that $|\nabla P| \ll|P \nabla \theta|$. When this condition is met, we say that the amplitude $P(\bar{x})$ is a slowly varying function of position. A further assumption comes from differentiation of Eq.(11.1.2), which gives

$$
\begin{equation*}
\nabla \theta=-\omega \nabla\left(\frac{\bar{x} \cdot \bar{e}(\bar{x})}{c(\bar{x})}\right)=-\frac{\omega \bar{e}}{c(\bar{x})}+\theta \frac{\nabla c}{c(\bar{x})} \tag{11.1.6}
\end{equation*}
$$

Unless the fluid is homogeneous, $\nabla c$ is not zero. A geometrical acoustic analysis assumes that the sound speed varies slowly relative to the local wavelength, in which case $\nabla c \approx \overline{0}$. The combination of both assumptions leads to the geometrical acoustics assumption of the pressure gradient as

[^48]\[

$$
\begin{equation*}
\nabla p \approx-\operatorname{Re}\left(i \frac{\omega \bar{e}}{c(\bar{x})} P e^{i(\omega t-\theta(\bar{x}))}\right) \tag{11.1.7}
\end{equation*}
$$

\]

Another requirement, which will be relaxed later, is that the radius of curvature of the wavefront should be large compared to the wavelength. When both conditions are met, the signal may be considered to be locally planar. Conversely, ray theory might require correction, or it might not be valid at all, if either property is not obtained in some region.

The extension of Eq. (11.1.1) to an arbitrary signal is

$$
\begin{equation*}
p(\bar{x}, t)=F\left(\bar{x}, t^{\prime}\right), \quad t^{\prime}=t-\tau(\bar{x}) \tag{11.1.8}
\end{equation*}
$$

The essence of the geometrical acoustics approximation is that the explicit dependence of $F$ on $\bar{x}$ is taken to be very gradual relative to the manner in which $\tau(\bar{x})$ varies spatially. This leads to the pressure gradient being represented by

$$
\begin{equation*}
\nabla p \approx-\frac{\partial F}{\partial t^{\prime}} \nabla \tau \tag{11.1.9}
\end{equation*}
$$

The nature of the wavefronts is defined by Eq.(11.1.4). An examination of Eqs. (11.1.1) and (11.1.8) indicates that knowledge of the wavefronts is fundamental to describing the pressure field. This observation should cause the reader to wonder how rays enter the picture? That question is addressed by Fig. 11.1, which follows a zero-phase wavefront. At instant $t$, a specific signal is located at $x=\bar{\xi}$. The location $\bar{\xi}$ shifts in time by the amount necessary to keep it on the wavefront. The locus of positions at which this signal is located is a curved line whose equation is $\bar{x}=\bar{\xi}(t)$. This is a parametric description of a ray, in which time is the parameter. Of course, at this stage we do not know the functional dependence.

Fig. 11.1 Relationship of rays and wavefronts in the geometrical acoustics approximation. The normal $\bar{e}$ to the wavefront at any location is the direction in which the signal at that location propagates


If we know the wavefront at any instant $t$ and the sound speed at any location, there is a graphical construction that will lead to wavefronts and rays at later times. We consider a single step from $t$ to $t+d t$. The planar nature of the propagation means that
the phase velocity at $\bar{\xi}$ is $c(\bar{\xi})$ in the direction that is normal to the wavefront at $\bar{\xi}$. The unit vector for this normal direction is $\bar{e}(\bar{\xi})$ in Fig. 11.1. Given the current wavefront, we can draw this normal, either manually or with a CAD program. In a time interval $d t$, the signal propagates through distance $c(\bar{\xi}) d t$, which we lay out along $\bar{e}(\bar{\xi})$. This gives the next point on the ray as $\bar{x}=\bar{\xi}(t+d t)=\bar{\xi}(t)+\bar{e}(\bar{\xi}) c(\bar{\xi}) d t$. If we apply this construction to many points on the wavefront at $t$, we will establish where the wavefront is at $t+d t$. Thus, we have updated the location of the wavefront. The process may be repeated as long as necessary. If we track the position of a particular point on the initial waveform as time evolves, we will obtain the ray path for the signal that was at that initial point at that initial instant.

There are several reasons why this graphical procedure is only used conceptually. Most significantly, it is approximate because it can only be implemented for a finite time step. In addition, any method we use to graphically identify $\bar{e}(\bar{\xi})$ will inherently be approximate. There is no way to quantify either type of error. Beyond considerations of accuracy, the procedure is tedious, because many time steps would be required to identify the rays and wavefronts over a reasonably large interval.

As we progress through this chapter, we will encounter two approaches by which geometrical acoustics analyses are implemented. The first, which is limited to fluids whose properties depend on a single spatial coordinate, directly determines the rays. The second approach is suitable for an arbitrary heterogeneity. It simultaneously determines wavefronts and rays. Both formulations are used primarily as the basis for computational algorithms that we will explore. However, before we can address either formulation, it is necessary that we identify the governing field equation.

### 11.1.1 Field Equations for an Inhomogeneous Fluid

The ambient density $\rho_{0}$ and pressure $p_{0}$ are known function of position, and the fluid is taken to be at rest in this state. We seek to determine the acoustic perturbations, $\rho^{\prime}(\bar{x}, t), p^{\prime}(\bar{x}, t)$, and $\bar{v}(\bar{x}, t)$. (It is convenient to temporarily revert to the use of a prime to denote increments of state variables from their ambient value.) A body force acting on the fluid may depend on the density, so it is denoted as $\bar{b}\left(\bar{x}, \rho_{0}+\rho^{\prime}\right)$. The body force adds a term to the momentum equation, Eq. (4.1.7), but the general continuity equation, Eq. (4.1.5), is unmodified.

A linearized set of equations is obtained by introducing the small signal approximations,

$$
\begin{equation*}
\rho=\rho_{0}(\bar{x})+\rho^{\prime}(\bar{x}, t), \quad p=p_{0}(\bar{x})+p^{\prime}(\bar{x}, t), \quad \bar{v}=\bar{v}^{\prime}(\bar{x}, t) \tag{11.1.10}
\end{equation*}
$$

in which $\left|\rho^{\prime}\right| \ll \rho_{0},\left|p^{\prime}\right| \ll p_{0}$, and $|\bar{v}|$ is much less than the speed of sound at ambient conditions. When we introduce this decomposition into the continuity equation and drop all terms that are products of primed quantities, we find that

$$
\begin{equation*}
\frac{\partial \rho^{\prime}}{\partial t}+\nabla \cdot\left(\rho_{0} \bar{v}^{\prime}\right)=0 \tag{11.1.11}
\end{equation*}
$$

Substitution of the small-signal representation into the momentum equation gives

$$
\begin{equation*}
\frac{\partial \bar{v}^{\prime}}{\partial t}=-\left(\frac{\nabla p_{0}+\nabla p^{\prime}}{\rho_{0}}\right)+\frac{1}{\rho_{0}}\left[\bar{b}\left(\bar{x}, \rho_{0}\right)+\left(\frac{\partial \bar{b}}{\partial \rho}\right)_{0} \rho^{\prime}\right] \tag{11.1.12}
\end{equation*}
$$

Note that we have used the smallness of $\rho^{\prime}$ to replace $\bar{b}$ by its first two terms in a Taylor series relative to the ambient state. All primed variables are zero in the ambient state, so the ambient pressure must satisfy

$$
\begin{equation*}
\nabla p_{0}=\bar{b}\left(\bar{x}, \rho_{0}\right) \tag{11.1.13}
\end{equation*}
$$

Hence, the linearized momentum equation is

$$
\begin{equation*}
\frac{\partial \bar{v}^{\prime}}{\partial t}=-\frac{\nabla p^{\prime}}{\rho_{0}}+\left(\frac{\partial \bar{b}}{\partial \rho}\right)_{0} \frac{\rho^{\prime}}{\rho_{0}} \tag{11.1.14}
\end{equation*}
$$

Three state variables appear in the continuity and momentum equations. The third relation is the equation of state. The assumption that heat flow is negligible remains valid. However, because the entropy now may be position-dependent, the pressure must be taken to depend on both the density and entropy, $p=p(\rho, S)$. Negligible heat flow between particles corresponds to taking the entropy of a specific particle to be constant as the particle moves through space. The material derivative describes how a property of a moving particle changes as it moves. Thus, we seek a relation between $p^{\prime}$ and $\rho^{\prime}$ based on $d S / d t=0$. The total pressure is $p=p_{0}+p^{\prime}$, but $d p_{0} / d t=0$ because the ambient pressure is taken to vary only with position. Thus, the material derivative of the pressure equation of state is

$$
\begin{equation*}
\frac{d p^{\prime}}{d t}=\frac{\partial p}{\partial \rho} \frac{d \rho}{d t}+\frac{\partial p}{\partial S} \frac{d S}{d t}=\frac{\partial p}{\partial \rho}\left(\frac{\partial \rho}{\partial t}+\bar{v}^{\prime} \cdot \nabla \rho\right) \tag{11.1.15}
\end{equation*}
$$

This expression does not invoke the small-signal approximation. To do so we introduce $\rho=\rho_{0}(\bar{x})+\rho^{\prime}$ and drop higher order terms. Furthermore, the small-signal approximation indicates that $d p^{\prime} / d t$ and $\partial p^{\prime} / \partial t$ differ by second-order terms, which are dropped. Thus, the equation of state for a heterogeneous fluid is a linear differential equation,

$$
\begin{equation*}
\frac{\partial p^{\prime}}{\partial t}=c^{2}\left(\frac{\partial \rho^{\prime}}{\partial t}+\bar{v}^{\prime} \cdot \nabla \rho_{0}\right), \quad c^{2}=\left(\frac{\partial p}{\partial \rho}\right)_{0} \tag{11.1.16}
\end{equation*}
$$

Although this definition of $c$ is the same as it is for a homogeneous fluid, $c$ now is a function of position.

We have derived three differential equations that govern $p^{\prime}, \rho^{\prime}$, and $\bar{v}^{\prime}$, but we seek a single differential equation for $p^{\prime}$. Toward that end, we substitute $\partial \rho^{\prime} / \partial t$ from

Eq. (11.1.11) into the equation of state, which leads to

$$
\begin{equation*}
\frac{\partial p^{\prime}}{\partial t}=-\rho_{0} c^{2} \nabla \cdot \bar{v}^{\prime} \tag{11.1.17}
\end{equation*}
$$

Next, we take the time derivative of this relation. Substitution of $\partial \bar{v} / \partial t$ from Eq. (11.1.14) removes the particle velocity. The result is

$$
\begin{equation*}
\frac{1}{\rho_{0} c^{2}} \frac{\partial^{2} p^{\prime}}{\partial t^{2}}=-\nabla \cdot \frac{\partial \bar{v}^{\prime}}{\partial t}=\nabla \cdot\left[\frac{\nabla p^{\prime}}{\rho_{0}}-\left(\frac{\partial \bar{b}}{\partial \rho}\right)_{0} \frac{\rho^{\prime}}{\rho_{0}}\right] \tag{11.1.18}
\end{equation*}
$$

A rearrangement of terms yields

$$
\begin{equation*}
\rho_{0} \nabla \cdot\left(\frac{\nabla p^{\prime}}{\rho_{0}}\right)-\frac{1}{c^{2}} \frac{\partial^{2} p^{\prime}}{\partial t^{2}}=\rho_{0} \nabla \cdot\left[\left(\frac{\partial \bar{b}}{\partial \rho}\right)_{0} \frac{\rho^{\prime}}{\rho_{0}}\right] \tag{11.1.19}
\end{equation*}
$$

If the term in the right side were not present, this would be a field equation for $p^{\prime}$. Indeed, most references do not list this term because they ignore the body force at the outset. To see why it is allowable to do so, consider a plane wave propagating in direction $\bar{e}$ in the ocean. The body force per unit volume is gravitational attraction, so $\bar{b}=\rho g \bar{e}_{z}$ and $(\partial \bar{b} / \partial \rho)_{0}=g \bar{e}_{z}$. Let us use the relations for a homogeneous fluid to compare the two terms whose divergence is taken in the above. Thus, we set $p^{\prime}=\operatorname{Re}(\hat{P} \exp (i \omega t-i k \bar{e} \cdot \bar{x}))$ and $\rho^{\prime} \approx p^{\prime} / c^{2}$. This leads to estimates of the two terms as

$$
\begin{equation*}
\left|\frac{\nabla p^{\prime}}{\rho_{0}}\right| \approx \frac{k\left|p^{\prime}\right|}{\rho_{0}}, \quad\left|\left(\frac{\partial \bar{b}}{\partial \rho}\right)_{0} \frac{\rho^{\prime}}{\rho_{0}}\right| \approx \frac{g\left|p^{\prime}\right|}{\rho_{0} c^{2}} \tag{11.1.20}
\end{equation*}
$$

The ratio of the second term to the first is $g /(\omega c)$, which is extremely small even at $\omega=1 \mathrm{rad} / \mathrm{s}$, unless the fluid is remarkably slow. For this reason, we shall proceed without considering the body force effect. (It might not be allowable to do in some situations, such as a strong magnetic field confining a plasma).

In the case of a gas, notably the atmosphere, dropping the body force term in Eq. (11.1.19) is the sole fundamental simplification of the field equation. The equation for a heterogeneous gas therefore is

$$
\begin{equation*}
\nabla^{2} p^{\prime}-\frac{1}{\rho_{0}(\bar{x})} \nabla \rho_{0} \cdot \nabla p^{\prime}-\frac{1}{c(\bar{x})^{2}} \frac{\partial^{2} p^{\prime}}{\partial t^{2}}=0 \tag{11.1.21}
\end{equation*}
$$

If the density changes gradually, the preceding field equation may be simplified further by setting $\nabla \rho_{0}$ to zero. To determine the conditions for which this approximation is reasonable, let us again make use of the description of a plane wave in a homogeneous fluid to estimate the effect. The density gradient is approximated
as the difference $\Delta \rho^{\prime}$ between the minimum and maximum values divided by the distance $\ell$ over which these extreme values are observed. For $\rho_{0}$, we use the average value. For a harmonic plane wave propagating in direction $\bar{e}$, we have $\nabla p^{\prime}=-i k \bar{e} p^{\prime}$. The maximum magnitude of $\nabla \rho_{0} \cdot \nabla p^{\prime}$ occurs if $\bar{e}$ is parallel to the density gradient. Thus, we have the estimates that

$$
\begin{equation*}
\nabla^{2} p^{\prime} \approx-k^{2} p^{\prime}, \frac{\nabla \rho_{0}}{\rho_{0}} \cdot \nabla p^{\prime} \approx-i k \frac{\Delta \rho^{\prime}}{\rho_{\mathrm{av}} \ell} p^{\prime} \tag{11.1.22}
\end{equation*}
$$

With $k$ replaced by $2 \pi / \lambda$, the order of magnitude of the ratio of these terms is $O\left(\left(\Delta \rho / \rho_{\mathrm{av}}\right) /(2 \pi \ell / \lambda)\right)$.

Density variations in the atmosphere can be quite large, but that is not true for water. To demonstrate this assertion consider standard formulas for the density of water provided by Pierce. ${ }^{5}$ The value of $\rho_{0}$ depends on the ambient pressure and temperature, and salinity in the case of the ocean. The reference value for seawater is $\rho_{0}=1027 \mathrm{~kg} / \mathrm{m}^{3}$ at $10^{\circ} \mathrm{C}$. Raising the ambient pressure to 2 MPa , which corresponds to a 200 m depth, raises the density by less than $0.1 \%$. A temperature increase to $30^{\circ} \mathrm{C}$ decreases the density by less than $2 \%$. Distilled water exhibits similar tendencies. The density of seawater also depends on salinity, whose reference value is 35 parts per thousand (ppt). It is very rare that the salinity varies by more than 2 ppt from the reference, and such a fluctuation changes the density by $0.15 \%$. Furthermore, it is unusual that these extreme changes of ambient conditions occur over less than a kilometer. In contrast, the wavelength at 10 Hz is approximately 150 m , and the wavelength decreases with increasing frequency. In other words, $2 \pi \ell / \lambda$ will be large. Thus, neglecting the contribution of the density gradient to the field equation is well warranted if the medium is a large body of water or most other liquids. One exception occurs at the interface between distinct layers, which was studied in Chap.5. The geometrical acoustic formulation addresses discontinuous changes of density and sound speed as separate considerations.

When the simplifications that ignore the effects of a body force and of a spatially varying density are invoked, what remains of the field equation is

$$
\begin{equation*}
\nabla^{2} p^{\prime}-\frac{1}{c(\bar{x})^{2}} \frac{\partial^{2} p^{\prime}}{\partial t^{2}}=0 \tag{11.1.23}
\end{equation*}
$$

This equation resembles the linear wave equation for a homogeneous fluid. However, the spatial dependence of $c$ has a fundamental effect.

It should be apparent at this juncture that any analysis we pursued in previous chapters probably will not be valid for a heterogeneous fluid because the field equation is a partial differential equation with variable coefficients. Indeed, any direct mathematical analysis would be challenging. Nevertheless, it is important that we have derived this equation because it validates considering the pressure to behave

[^49]locally like a plane wave. In addition, Eq.(11.1.21) will be crucial for the analysis of the pressure amplitude in an fluid whose heterogeneity is an arbitrary function of position.

### 11.1.2 Reflection and Refraction of Rays

The geometrical acoustic approximation that a ray behaves locally as though it were planar has some important corollaries. Suppose a wave is incident on a locally reacting solid surface. The reflected wave may be identified by adapting the laws for reflection of a plane wave at oblique incidence to a planar boundary. In keeping with the restrictions of geometrical acoustics, the surface's principal radii of curvature are required to be much greater than a typical wavelength. When this restriction is fulfilled, the surface may be replaced by the tangent plane at the point of incidence. Other than the dependence of the normal $\bar{n}$ on the point of incidence, the reflection process is the same as that of planar waves from a flat surface. Thus, the normal $\bar{n}$ and the incident ray form a plane. The reflected ray lies in this plane at the same angle as the angle of incidence. A consequence of this behavior is that rays that were parallel prior to reflection will not be parallel after reflection, as is illustrated by Fig.11.2.

Fig. 11.2 Reflection of rays from a curved surface


At very high frequencies, surfaces that fit a local impedance model tend to be very stiff. For that reason, solid surfaces are typically considered to be rigid in the geometrical acoustics approximation. Then, the amplitude of the reflected wave at the point of incidence equals that of the incident wave. If the impedance $Z$ is finite, the reflection coefficient in Eq. $(5.2 .15)$ may be used,

$$
\begin{equation*}
P_{\mathrm{R}}=R P_{\mathrm{I}}, \quad R=\frac{Z-\rho_{0} c / \cos \theta}{Z+\rho_{0} c / \cos \theta} \tag{11.1.24}
\end{equation*}
$$

As is true for plane waves, a description using a reflection coefficient should be limited to frequency-domain analyses, unless $Z$ is real.

It is unusual that two fluids meet at a curved interface. A possibility is that two fluids are separated by a highly compliant solid sheet. Like the case of reflection from a solid surface, the reflection and transmission of an incident ray is defined relative
to the normal at the point of incidence. The process is illustrated in Fig. 11.3. The reflected and refracted rays lie in the plane formed by the incident ray and the normal at the point of incidence. The angle of reflection equals the angle of incidence. The angle of transmission is governed by Snell's law, which states that

$$
\begin{equation*}
\frac{\sin \theta_{2}}{c_{2}}=\frac{\sin \theta_{1}}{c_{1}} \tag{11.1.25}
\end{equation*}
$$

Fig. 11.3 Transmission and reflection of a ray incident on the curved interface of two fluids


The amplitudes of the signal on the reflected and transmitted rays are taken to be those derived for a plane wave. The reflection and transmission coefficients are as given in Eq. (5.3.8),

$$
\begin{equation*}
R=\frac{Z_{2}-Z_{1}}{Z_{2}+Z_{1}}, \quad T=\frac{2 Z_{2}}{Z_{2}+Z_{1}}, \quad Z_{j}=\frac{\rho_{j} c_{j}}{\cos \theta_{j}} \tag{11.1.26}
\end{equation*}
$$

The angle of transmission is measured relative to the local normal. Consequently, even if incident rays are parallel, the amplitudes of transmitted and reflected rays will vary with position on the interface.

We saw in Sect. 5.3 that supercritical incidence occurs if $\cos \theta_{1}>c_{1} / c_{2}$, which obviously can only occur if $c_{2}>c_{1}$. Supercritical incidence leads to a wave in the receiving fluid that evanesces with increasing distance perpendicularly the interface. In the case of a truly planar interface, such a wave propagates parallel to the surface. In a strict sense such a wave does not fit the geometrical acoustic model.

EXAMPLE 11.1 A plane wave propagates parallel to the axis of symmetry of the cap of a sphere whose radius is $a$. The wave is reflected from the inner concave surface of the cap, which is rigid. Determine and graph the paths of the reflected rays.

## Significance

This system, which is the acoustical analog of the mirror in a reflecting optical telescope, has an acoustic application as part of an extremely sensitive microphone. The only mathematical tools required for the analysis are those of trigonometry, so the basic aspects of ray paths will be easy to identify. The results will exhibit the phenomenon of a caustic, at which rays are concentrated.

## Solution

The system is axisymmetric, so the picture is the same for any slicing plane that contains the axis of the spherical cap. The range of polar angles $\psi \geq 0$ for one slice is depicted in Fig. 1. The spherical radius of the cap is $a$. The sketch shows a ray at an arbitrary distance $R$ from the axis of symmetry. This defines the polar angle $\psi$ to the point where the ray is incident on the mirror according to $\psi=\sin ^{-1}(R / a)$. The normal direction $\bar{n}$ coincides with the radial line. Because the angle of reflection is $\psi$, the reflected ray, the radial line, and the axis of symmetry form an isosceles triangle. The distance where the reflected ray intersects the axis of symmetry is $s=(a / 2) / \cos \psi$ inward from the center.


Figure 1.

A full picture requires consideration of many rays, so we shall consider a range of $R$ values. If the maximum $\psi$ is too large, a reflected ray will go on to be incident at a second point on the cap. This situation may be avoided if the cap is such that the maximum polar angle is less than $45^{\circ}$. We will consider a cap whose maximum polar angle is $30^{\circ}$.

Construction of the rays is assisted by defining a coordinate system whose $x$-axis coincides with the axis of symmetry, with the origin at the center. We construct a set of incident rays, spaced at equal increments of the offset distance. Because of the axisymmetric nature of the system, it is sufficient to consider positive polar angles. For each $R$, we draw an incident ray from $(0, R)$ to $(a \cos \psi, R)$, and a reflected ray from $(a \cos \psi, R)$ to ( $s, 0$ ), where $s$ is described above. In Fig. 2, the reflected ray are extended beyond the axis of symmetry.


Figure 2.
Every cutting plane features the same picture. Furthermore, Fig. 2 only shows half the rays in a cutting plane because there are additional rays for $\psi<0$. The full set for a cross section are depicted in Fig. 3. The reflected rays extend beyond the window that is plotted. The collection of rays corresponding to all cutting planes represents an axisymmetric distribution. Thus, the locus of straight rays at offset distance $R$ is a cylinder, and the locus of reflected rays corresponding to a specified $R$ is a cone whose apex is situated at distance $s$.


Figure 3.

There is a region near the $z$-axis where the rays are closely spaced. Our later investigations will show that the amplitude grows in regions where the spacing of rays decreases. Another interesting aspect is this region is delimited by a curve beyond which there are no neighboring reflected rays. To better understand this feature Fig. 4 zooms in on the reflected rays in this region of concentration. The decreased spacing between rays and the absence of rays beyond a bounding curve are produced by nearly parallel rays intersecting. The minimum $s$ from such an intersection to the center of the spherical cap is the ray associated with $R \rightarrow 0$. This corresponds to $\psi \rightarrow 0$, for which $s \rightarrow a / 2$. The bounding curve generated by the intersecting rays is called a caustic, and $s=a / 2$ marks the arrête, which is the term given to the location where a caustic terminates.


Figure 4.

In some applications, it is desirable that all rays intersect in a very compact region. Convergence at a point, which is the definition of a focus, can be attained with a mirror that is the interior of a paraboloid of revolution. (This is the subject of Homework Problem 11.1.) All rays for the case of a spherical cap nearly converge at a point if the largest polar angle is relatively small. However, there is no size that leads to a perfect focus. In optics, this is known as spherical aberration.

### 11.2 Propagation in a Vertically Stratified Medium

An important model in environmental acoustics considers the fluid to have properties that vary with vertical elevation, which is designated as $z$. The ambient density, pressure, and temperature are taken to depend solely on $z$, and the fluid is quiescent in the ambient state. The analysis that follows parallels the manner in which a classic textbook on underwater acoustics ${ }^{6}$ begins. The development has the attractive feature of emphasizing the physical properties of rays. Modifications to account for flow

[^50]in the ambient state, such as winds in the atmosphere, are best deferred until the basic concepts of geometrical acoustics are understood. The text by Pierce ${ }^{7}$ offers an excellent portal to this subject. For underwater acoustics flow is less of an issue because the flow rates are lower than in the atmosphere, and the sound speed is much higher.

### 11.2.1 Snell's Law for Vertical Heterogeneity

We begin with the case where finite thickness layers are stacked vertically. Propagation in the ocean typically is analyzed by measuring $z$ downward with $z=0$ at the free surface. That is the convention used here. A ray starting in layer $n$ behaves like a plane wave, so it is transmitted into layer $n+1$. The direction of the transmitted ray is governed by Snell's law. Consequently, the ray will remain in a constant vertical plane. Whereas the treatment of transmission in Sect. 5.3 used the angle of incidence, the developments that follow use the grazing angle $\chi$ of the ray, which is measured from the horizontal to the ray. The arrangement is depicted in Fig. 11.4. It will be necessary to distinguish between propagation that is downward (increas$\operatorname{ing} z$ ) and upward (decreasing $z$ ). This can be done by letting $\chi$ be in the range $-\pi / 2 \leq \chi \leq \pi / 2$, with negative values corresponding to upward propagation.

Fig. 11.4 Definition of the grazing angle for propagation in a vertically stratified fluid


Let "I" denote the conditions at the location where the ray was generated. Then, Snell's law requires that

$$
\begin{equation*}
\frac{c_{n+1}}{\cos \chi_{n+1}}=\frac{c_{n}}{\cos \chi_{n}}=\cdots=\frac{c_{\mathrm{I}}}{\cos \chi_{\mathrm{I}}} \tag{11.2.1}
\end{equation*}
$$

The situation where the sound speed $c(z)$ is a continuous function corresponds to layers having differentially small thickness. This condition is described by replacing $c_{n}$ with $c(z)$, and $\chi_{n}$ with $\chi(z)$. Thus, the grazing angle at any depth is defined by

$$
\begin{equation*}
\cos \chi=\frac{1}{\mu(z)} \cos \chi_{\mathrm{I}}, \quad \mu=\frac{c_{\mathrm{I}}}{c(z)} \tag{11.2.2}
\end{equation*}
$$

The quantity $\mu$ is the index of refraction. It captures the manner in which the depth dependence affects any ray, whereas the $\cos \chi_{\mathrm{I}}$ factor is specific to one ray. It is

[^51]evident that the grazing angle changes continuously if $c$ is a continuous function of depth. This means that each ray path will be curved, so that $\chi$ measures the angle between the tangent to this curve and the horizontal.

It is worth noting that no consideration has been given to a reflected ray. The reflection coefficient depends on the difference of the $\rho c / \sin \chi$ values on either side of the interface. A differential difference of these value associated with a continuous variation of $\rho$ and $c$ leads to a reflection coefficient that is infinitesimal.

The observation that $\chi$ defines the local tangent to the ray path is the basis for identification of these paths. Consider Fig. 11.5, which depicts a differential arc $d \ell$. Its slope is $d z / d x$, so we have

$$
\begin{equation*}
d x=\frac{d z}{\tan \chi}=\frac{\cos \chi}{\left[1-(\cos \chi)^{2}\right]^{1 / 2}} d z, \quad d \ell=\frac{d z}{\sin \chi}=\frac{1}{\left[1-(\cos \chi)^{2}\right]^{1 / 2}} d z \tag{11.2.3}
\end{equation*}
$$

Snell's law and trigonometry lead to descriptions of both differential distances in terms of $\cos \chi_{\mathrm{I}}$,

$$
\begin{align*}
& d x=\frac{\cos \chi_{\mathrm{I}}}{\left[\mu^{2}-\left(\cos \chi_{\mathrm{I}}\right)^{2}\right]^{1 / 2}} d z  \tag{11.2.4}\\
& d \ell=\frac{\mu}{\left[\mu^{2}-\left(\cos \chi_{\mathrm{I}}\right)^{2}\right]^{1 / 2}} d z
\end{align*}
$$

Fig. 11.5 Pythagorean theorem for the arclength along a ray


The first of Eq.(11.2.4) defines $x$ as a function of $z$. Because $\mu$ is a function of $z$, the relation $x(z)$ may be found as a definite integral starting from the first location $\left(x_{\mathrm{I}}, z_{\mathrm{I}}\right)$. At the location where the ray is generated, the sense of the ray might be downward. In that case $z>z_{\mathrm{I}}$. Upward propagation leads to $z<z_{\mathrm{I}}$. In either case, $x$ increases monotonically. Both cases are captured by taking the absolute value of the integral, so that

$$
\begin{equation*}
x=x_{\mathrm{I}}+\left|\int_{z_{\mathrm{I}}}^{z} \frac{\cos \chi_{\mathrm{I}}}{\left[\mu(\xi)^{2}-\left(\cos \chi_{\mathrm{I}}\right)^{2}\right]^{1 / 2}} d \xi\right| \tag{11.2.5}
\end{equation*}
$$

The second of Eq. (11.2.4) is useful because the local propagation speed in geometrical acoustics is $c(z)$. Thus, the time required for a signal to propagate distance $d \ell$ along a ray is $d t=d \ell / c(z) \equiv \mu d \ell / c_{\mathrm{I}}$. The arclength is described by the second of Eq.(11.2.4), so

$$
\begin{equation*}
d t=\frac{\mu(z)}{c_{\mathrm{I}}} d \ell=\frac{\mu(z)^{2}}{c_{\mathrm{I}}\left[\mu^{2}-\left(\cos \chi_{\mathrm{I}}\right)^{2}\right]^{1 / 2}} d z \tag{11.2.6}
\end{equation*}
$$

If a signal departs from the first location at $t_{\mathrm{I}}$, the time at which it arrives at another point on a specified ray is

$$
\begin{equation*}
t=t_{\mathrm{I}}+\left|\int_{z_{\mathrm{I}}}^{z} \frac{\mu(\xi)^{2}}{c_{\mathrm{I}}\left[\mu(\xi)^{2}-\left(\cos \chi_{\mathrm{I}}\right)^{2}\right]^{1 / 2}} d \xi\right| \tag{11.2.7}
\end{equation*}
$$

The absolute value of the integrand is used here because the elapsed time increases monotonically, regardless of whether the propagation is upward or downward.

Evaluation of Eq. (11.2.5) for a range of $z$ gives the ray path, and evaluation of Eq. (11.2.7) gives the time at which the signal will arrive at a point on that path. Propagation into a region where the sound speed decreases in the direction of propagation leads to increasing values of $|\chi|$. This is the situation depicted Fig. 11.6a and d. There is no limit to the vertical propagation in these cases, other than incidence at the bottom. The behavior of rays in cases where the sound speed increases in the direction of propagation are depicted in Fig. 11.6b and c. There the grazing angle progressively decreases.

Fig. 11.6 Influence of the gradient of the sound speed on the curvature of a ray

(c)



(d)

A simple explanation of the tendency for a ray to bend may be gained by considering a wave that initially is planar, with a wavefront that lies in the vertical plane. If the sound speed increases with depth, so that the signal propagates faster as the depth increases, the wavefront will advance progressively farther as the depth increases. The direction of propagation is perpendicular to the wavefront, therefore the ray begins to point upward. The case where $d c / d z<0$ leads to the opposite effect, wherein the wavefront advances more as the depth decreases.

It might ultimately happen that the grazing angle is zero. The sound speed at which this condition occurs will be denoted as $c_{0}$. Setting $\chi=0$ in Snell's law gives

$$
\begin{equation*}
c_{0}=\frac{c_{\mathrm{I}}}{\cos \chi_{\mathrm{I}}} \tag{11.2.8}
\end{equation*}
$$

This value is dependent on the initial grazing angle. If a ray reaches a depth at which $c(z)=c_{0}$ for its associated $\chi_{\mathrm{I}}$, then it will attain a horizontal tangency, which is the limit of its vertical propagation. This condition occurs only if there is a depth at which $c(z)$ equals the value of $c_{0}$ for the ray. For instance, rays that are launched vertically remain straight. After a ray attains this horizontal tangency, the vertical sense of the propagation reverses, so the horizontal tangency is said to be a turning point.

It might happen that there are turning points above and below the launch point. In Fig. 11.7a $c\left(z_{B}\right)$ and $c\left(z_{D}\right)$ equal $c\left(z_{A}\right) / \cos \chi_{\mathrm{I}}$. The ray is launched downward, but it curves upward because $d c / d z>0$. Beyond turning point $B$ it continues to curve upward until it reaches a depth at which $d c / d z<0$. This results in downward curvature. Upward propagation ceases at turning point $D$, and beyond that point the propagation direction is downward. It follows that that this ray is confined to the range $z_{D} \leq z \leq z_{B}$. Depending on the sound speed profile $c(z)$ and the value of $\chi_{\mathrm{I}}$, other rays will be confined differently or not at all.

Fig. 11.7 Turning points and the surface limit the vertical propagation of rays


An exceptional situation occurs if a turning point occurs at the depth where the sound speed is a maximum. In the region above that depth, a ray curves upward, whereas a ray in the region below that depth curves downward. At the turning point the propagation direction is horizontal. If $c$ is a maximum at that depth the tendency for the horizontal ray to curve downward is equal to the tendency for it to curve upward. The consequence is that the horizontal ray splits in two, with one part going up and the other going down.

Another possible condition is that a ray arrives at the free surface or the horizontal bottom of a channel. Such a ray is reflected from the boundary. Because of the locally planar nature of the signal, the reflected ray obeys the law for reflection of plane waves at a free surface. Thus, $\chi$ for the reflected ray is $-\chi$ for the incident one. This is the situation in Fig. 11.7b. The launch angle $\chi_{\mathrm{I}}$ is larger, so the lower turning point occurs
at a greater depth $z_{B^{\prime}}$. Reflection at the surface will occur if there is no turning point at a lesser depth than $z_{\mathrm{I}}$, or if the depth at which $c(z)=c_{\mathrm{I}} / \cos \chi_{\mathrm{I}}$ is negative.

An interesting feature of a ray path results from the observation that $|\chi|$ for a specific ray depends only on $z$. Thus, the grazing angle at a given depth beyond a turning point is the negative of the value at that depth prior to the turning point. This means that a ray is symmetric with respect to a vertical line through a turning point. This feature is exhibited in Fig. 11.7 which shows that the grazing angle at the original depth $z_{A}$ is either $\chi_{\mathrm{I}}$ or $-\chi_{\mathrm{I}}$, depending on whether the propagation is downward or upward. This attribute also applies relative to a surface reflection point. Furthermore, the situation at points $E$ and $E^{\prime}$ in Fig. 11.7 is the same as point $A$, except that the horizontal range is different. Thus, the pattern beyond these points is a periodic repetition of the pattern from point $A$.

It is useful to consider how to handle the integrals in Eqs. (11.2.5) and (11.2.7) in view of the nature of the ray paths in Fig. 11.7. From Snell's law, we know that $0 \leq\left|\cos \chi_{\mathrm{I}}\right| \leq \mu(z)$. From the starting location to the lower turning point, where $z=z_{B}$ and $x=x_{B}$, the integral's upper limit progressively increases. Consequently, the integral leads to a value of $x$ that increases monotonically with increasing $z$. Beyond the turning point, the upper limit decreases. The consequence is that both integrals decrease from their values at the turning point. This cannot be correct because it leads to $x$ values beyond the turning point that are less than $x_{B}$, and elapsed times required for the signal to propagate beyond the turning point that are less than the time to reach the turning point. This seeming paradox is actually a consequence of ignoring the existence of a singularity. The condition for a turning point is $\mu(z)=\cos \chi_{\mathrm{I}}$, so the integrands are singular at these points. To go beyond this point, we must perform piecewise evaluation of the integrals. That is, if $x_{\mathrm{tp}}, z_{\mathrm{tp}}$ are the coordinates of a turning point, then we must perform the evaluations as

$$
\begin{array}{r}
x=x_{\mathrm{I}}+\left|\int_{z_{\mathrm{I}}}^{z_{\mathrm{tp}}} \frac{\cos \chi_{\mathrm{I}}}{\left[\mu(\xi)^{2}-\left(\cos \chi_{\mathrm{I}}\right)^{2}\right]^{1 / 2}} d \xi\right|+\left|\int_{z_{\mathrm{tp}}}^{z} \frac{\cos \chi_{\mathrm{I}}}{\left[\mu(\xi)^{2}-\left(\cos \chi_{\mathrm{I}}\right)^{2}\right]^{1 / 2}} d \xi\right| \\
t=t_{\mathrm{I}}+\left|\int_{z_{\mathrm{I}}}^{z_{\mathrm{tp}}} \frac{\mu(\xi)^{2}}{c_{\mathrm{I}}\left[\mu(\xi)^{2}-\left(\cos \chi_{\mathrm{I}}\right)^{2}\right]^{1 / 2}} d \xi\right|+\left|\int_{z_{\mathrm{tp}}}^{z} \frac{\mu(\xi)^{2}}{c_{\mathrm{I}}\left[\mu(\xi)^{2}-\left(\cos \chi_{\mathrm{I}}\right)^{2}\right]^{1 / 2}} d \xi\right| \tag{11.2.9}
\end{array}
$$

A better alternative is to recognize that the first term in the right side of these expressions gives $x_{\mathrm{tp}}$ and $t_{\mathrm{tp}}$. This allows us to write these equations as

$$
\begin{align*}
& x=x_{\mathrm{tp}}+\left|\int_{z_{\mathrm{tp}}}^{z} \frac{\cos \chi_{\mathrm{I}}}{\left.z \mu(\xi)^{2}-\left(\cos \chi_{\mathrm{I}}\right)^{2}\right]^{1 / 2}} d \xi\right| \\
& t=t_{\mathrm{tp}}+\left|\int_{z_{\mathrm{tp}}}^{z} \frac{\mu(\xi)^{2}}{c_{\mathrm{I}}\left[\mu(\xi)^{2}-\left(\cos \chi_{\mathrm{I}}\right)^{2}\right]^{1 / 2}} d \xi\right| \tag{11.2.10}
\end{align*}
$$

The same modifications are required to evaluate the ray path and elapsed time beyond a reflection point. The integrand is not singular at that point. Thus, the
integration may be carried without special consideration up to the reflection point, then using the values of $x$ and $t$ at the reflection point to proceed onwards. For example, in Fig. 11.7b, we may use the above procedure to find $x$ and $t$ for any point on segments $A^{\prime} B^{\prime}$ and $B^{\prime} C^{\prime} D^{\prime}$. To go beyond the reflection point, and thereby find segment $D^{\prime} E^{\prime}$, we would start with $x_{D^{\prime}}$ and $t_{D^{\prime}}$ and integrate over $0<z \leq z_{E^{\prime}}$. Beyond point $E^{\prime}$, periodic replication applies. The symmetry property of the waveform relative to turning points and waveforms could be used to reduce the integrations. In any event, this piecewise method of handling slope reversals at turning points and reflections is troublesome to incorporate into an autonomous computer program. An alternative approach is developed in the next example. It evaluates the integrals in a stepwise manner.

The curvature of rays is related to the gradient of the sound speed. To identify this relation, we again consider a differential arc $d \ell$, but now account for the change of slope along it. Thus, at the upper end of the element in Fig. 11.8 the grazing angle is $\chi$, while at the lower end it is $\chi+d \chi$. The center of curvature $C$ is the point at which lines perpendicular to the tangents intersect. The radius of curvature $\mathcal{R}$ is the distance from the center of curvature to a point on the arc at the specified $z$.

Fig. 11.8 Geometric construction of the radius of curvature $\mathcal{R}$ using the properties of a ray


The angle between these perpendiculars is the difference of the grazing angles, so the differential arclength is $d \ell=\mathcal{R} d \chi$. The grazing angle on a specific ray may be considered to be a function of the arclength, so $d \chi=(\partial \chi / \partial \ell) d \ell$. Implicit differentiation of Snell's law, Eq. (11.2.2), gives

$$
\begin{equation*}
-(\sin \chi) \frac{\partial \chi}{\partial \ell}=\frac{\partial}{\partial \ell}\left(\frac{\cos \chi_{\mathrm{I}}}{c_{I}} c(z)\right)=\frac{\cos \chi_{\mathrm{I}}}{c_{\mathrm{I}}}\left(\frac{d c}{d z} \frac{d z}{d \ell}\right) \tag{11.2.11}
\end{equation*}
$$

We found from Fig. 11.5 that $d z=(\sin \chi) d \ell$. Assembling these relations leads to

$$
\begin{equation*}
\frac{1}{\mathcal{R}}=\frac{\partial \chi}{\partial \ell}=-\frac{\cos \chi_{\mathrm{I}}}{c_{\mathrm{I}}} \frac{d c}{d z} \tag{11.2.12}
\end{equation*}
$$

Most treatments regard the radius of curvature as a positive number, so they take the absolute value of the preceding relation. This is not done here because the sign of $\mathcal{R}$ describes the direction in which the arc is curving. As shown in Fig. 11.6a and b , increasing $z$ leads to decreasing $c$. Therefore, the slope increases with depth, $d \chi / d z>0$ and $\mathcal{R}>0$. In Fig. 11.6c and d, the value of $c$ increases with increasing $z$. In that case, $d \chi / d z>0$ and $\mathcal{R}>0$. A different view is that the center of curvature lies below the ray path if $\mathcal{R}>0$, whereas it lies above the ray path if $\mathcal{R}<0$.


Fig. 11.9 Construction of circular ray paths for a fluid whose sound speed varies linearly with depth. The reference depth $z=0$ need not be situated in the fluid

Knowledge of the radius of curvature is most useful when the sound speed depends linearly on $z$, so that $d c / d z$ is constant. It is evident that $\mathcal{R}$ is constant in this case, but more information is embedded in Eq.(11.2.12). Let us define $z=0$ to be the depth at which $c=0$, so that $c=\beta z$. (It is not important that this reference depth probably is outside the fluid.) Correspondingly, we have

$$
\begin{equation*}
\mathcal{R}=-\frac{c_{\mathrm{I}}}{\left(\cos \chi_{\mathrm{I}}\right) \beta}=-\frac{c}{(\cos \chi) \beta}=-\frac{z}{\cos \chi} \tag{11.2.13}
\end{equation*}
$$

The center of curvature is at distance $\mathcal{R}$ in the direction of the inward normal to the ray. This distance is constructed in Fig. 11.9 for cases where $d c / d z$ is positive, $\beta>0$, and negative, $\beta<0$. In either case, the radius of curvature is the hypotenuse of a right triangle whose vertical side is $z$. The means that the center of curvature is situated at the virtual depth where the sound speed vanishes. This knowledge may be used for a graphical construction. Given the launch location and the initial grazing angle, a line perpendicular to the initial ray tangent is constructed. The intersection of this line with the $z=0$ reference elevation marks the fixed center of curvature. The ray path is a circle of radius $\mathcal{R}$ centered on this point.

It has been suggested that a case of arbitrary sound speed dependence may be analyzed by approximating it as a sequence of layers in which $c(z)$ depends linearly on $z$. This is equivalent to replacing $c(z)$ with a pointwise linear interpolation. Mathematical software enables direct evaluation of the integrals in Eqs.(11.2.5) and (11.2.7), so there is little reason to pursue an approximate procedure.

EXAMPLE 11.2 The sound speed profile in a certain region of the ocean varies parabolically with depth according to $c(z)=c\left(z_{\text {min }}\right)+$ $\left[c(0)-c\left(z_{\text {min }}\right)\right]\left(z / z_{\text {min }}-1\right)^{2}$, with $z_{\min }=200 \mathrm{~m}, c\left(z_{\text {min }}\right)=1460 \mathrm{~m} / \mathrm{s}$, and $c(0)=1520 \mathrm{~m} / \mathrm{s}$. Consider two rays emitted by a source at a depth of 150 m , both of which are launched toward the surface. One departs at a grazing angle of $12^{\circ}$, whereas the other is such that it has a turning point at the surface.

Determine and plot the path and the travel time along each ray as a function of the horizontal range $x$.

## Significance

The numerical algorithm developed herein handles turning points without the requirement of user intervention. The results will shed some light on the overall significance of sound speed variation on propagation in the ocean.

## Solution

The first task is to determine the range of $z$ covered by each ray. For the first, the initial grazing angle is $\chi_{\mathrm{I}}=-12^{\circ}$ and the sound speed at $z_{\mathrm{I}}=150 \mathrm{~m}$ is $c_{\mathrm{I}}=1463.8 \mathrm{~m} / \mathrm{s}$. The corresponding sound speed for a turning point is $c_{0} \equiv$ $c_{\mathrm{I}} / \cos \chi_{\mathrm{I}}=1496.5 \mathrm{~m} / \mathrm{s}$. There are two depths at which $c(z)=c_{0}: z_{B}=44.11 \mathrm{~m}$ and $z_{D}=355.87 \mathrm{~m}$. The initial grazing angle is upward, opposite the case in Fig. 11.7a. Thus, the ray first travels upward from $z_{\mathrm{I}}$ to $z_{B}$, then downward from $z_{B}$ to $z_{D}$, then upward from $z_{D}$ to $z_{B}$, and so on. We only need to follow one cycle, after which we may invoke the periodicity property. To determine the initial angle for the second ray, we compute $c$ at $z=0$, which is $1520 \mathrm{~m} / \mathrm{s}$. The grazing angle at this depth is zero in order to have a turning point. Snell's law states that the initial angle should be $\chi_{\mathrm{I}}=\cos ^{-1}\left(c(0) / c_{\mathrm{I}}\right)=-15.636^{\circ}$. The depth of the lower turning point corresponding to these values is 400 m .

Determination of a path requires evaluation of Eqs. (11.2.5) and (11.2.7) at a sequence of depths $z_{n}$. The computational effort will be reduced greatly if the integrals in Eqs. (11.2.5) and (11.2.7) are carried out in an incremental manner from one depth to the next. If the values of $x_{n-1}$ and $t_{n-1}$ at a specific $x_{n-1}$ have been determined, and the values at the next point are given by

$$
\begin{align*}
& x_{n}=x_{n-1}+\left|\int_{z_{n-1}}^{z_{n}} \frac{\cos \chi_{\mathrm{I}}}{\left[\mu(\xi)^{2}-\left(\cos \chi_{\mathrm{I}}\right)^{2}\right]^{1 / 2}} d \xi\right| \\
& t_{n}=t_{n-1}+\left|\int_{z_{n-1}}^{z_{n}} \frac{\mu(\xi)^{2}}{c_{\mathrm{I}}\left[\mu(\xi)^{2}-\left(\cos \chi_{\mathrm{I}}\right)^{2}\right]^{1 / 2}} d \xi\right| \tag{1}
\end{align*}
$$

In most cases, the integrals may be evaluated with standard software. Indeed, if $z_{n}-z_{n-1}$ is extremely small, a crude strip rule would be adequate. However, turning points are such that at some depth $x_{j}$ it happens that $\mu\left(x_{j}\right)=\cos \left(\chi_{\mathrm{I}}\right)$, so both integrands are singular at such points. This condition may be addressed by using Eq. (1) only if the value of $z$ at a turning point does not fall in the interval between $z_{n-1}$ and $z_{n}$. If it does, an alternative evaluation based on a series expansion of the denominator will allow us to pass the turning point. To see how to proceed, consider an upper turning point. Let $J$ be the index of this point, so that $z_{J}=z_{B}$. Because this is an upper point, adjacent points are at a greater depth. Thus, $z_{J-1}=z_{J+1}=z_{J}+\Delta$,
with $\Delta \ll 1$. The integration variable is expressed as $\xi=z_{B}+\varepsilon$. Series expansions based on the smallness of $\varepsilon$ in this interval give

$$
\begin{aligned}
c(z)= & c\left(z_{J}+\varepsilon\right) \approx c\left(z_{J}\right)+\varepsilon\left(\frac{d c}{d z}\right)_{J} \\
{\left[\mu(\xi)^{2}-\left(\cos \chi_{\mathrm{I}}\right)^{2}\right]^{-1 / 2} } & =\left[\frac{c_{\mathrm{I}}^{2}}{c\left(z_{J}+\varepsilon\right)^{2}}-\left(\cos \chi_{\mathrm{I}}\right)^{2}\right]^{-1 / 2} \\
& \approx\left[\frac{c_{\mathrm{I}}^{2}}{c\left(z_{J}\right)^{2}}\left[1-\frac{2 \varepsilon}{c\left(z_{J}\right)}\left(\frac{d c}{d z}\right)_{J}\right]-\left(\cos \chi_{\mathrm{I}}\right)^{2}\right]^{-1 / 2}
\end{aligned}
$$

Because $z_{J}$ is a turning point, it must be $c_{\mathrm{I}} / c\left(z_{J}\right)=\cos \chi_{\mathrm{I}}$. This observation leads to

$$
\begin{equation*}
\left[\mu(\xi)^{2}-\left(\cos \chi_{\mathrm{I}}\right)^{2}\right]^{-1 / 2} \approx\left[-\frac{c\left(z_{J}\right)^{3}}{2 c_{\mathrm{I}}^{2}\left(\frac{d c}{d z}\right)_{J}}\right]^{1 / 2} \varepsilon^{-1 / 2} \tag{11.2.14}
\end{equation*}
$$

Note that $d c / d z$ is negative at the upper turning point, so the square root yields a real value. Substitution of this approximation and $d \xi=d \varepsilon$ into Eq. (1) leads to

$$
\begin{align*}
& x_{J}-x_{J-1}=x_{J+1}-x_{J}=(2 \Delta)^{1 / 2}\left[\frac{c\left(z_{J}\right)}{\left|\left(\frac{d c}{d z}\right)_{J}\right|}\right]^{1 / 2} \\
& t_{J}-t_{J-1}=t_{J+1}-t_{J}=\frac{(2 \Delta)^{1 / 2}}{c_{I} \mu\left(z_{J}\right)}\left[\frac{c\left(z_{J}\right)}{\left|\left(\frac{d c}{d z}\right)_{J}\right|}\right]^{1 / 2} \tag{2}
\end{align*}
$$

The values of $x_{J-1}$ and $t_{J-1}$ presumably have been determined by invoking Eq. (1). The above expressions yield the corresponding values of $x_{J}$ and $t_{J}$, then $x_{J+1}$ and $t_{J+1}$. These expressions also apply to the case of a lower turning point, for which $(d c / d z)_{J}$ $>0$. In that case, the points neighboring the turning point will be at a slightly smaller depth, so that $z_{j-1}=z_{J+1}=z_{j}-\Delta$.

The actual ray path has two turning points. Figure 1 describes the scheme by which the depth $z$ is discretized. It is based on the recognition that a selected sampling interval $\Delta_{\text {I }}$ between the launch point and the first turning point might not fit the interval between the $z_{\mathrm{I}}$ and the lower turning point. Therefore, two intervals are used: $\Delta_{1}=\left(z_{\mathrm{I}}-z_{B}\right) / N_{1}$ for points at $z \leq z_{\mathrm{I}}$ and $\Delta_{2}=\left(z_{D}-z_{\mathrm{I}}\right) / N_{2}$ for points at $z \geq z_{\mathrm{I}}$. This leads to assignment of the depths according to $z_{n}=z_{\mathrm{I}}-(n-1) \Delta_{1}$ for $n=1$ to $N_{1}+1 ; z_{n}=z_{B}+\left(n-N_{1}-1\right) \Delta_{1}$ for $n=N_{1}+2$ to $2 N_{1}+1$; $z_{n}=z_{\mathrm{I}}+\left(n-2 N_{1}-1\right) \Delta_{2}$ for $n=2 N_{1}+2$ to $2 N_{1}+N_{2}+1$; and $z_{n}=z_{C}-$ $\left(n-2 N_{1}-N_{2}-1\right) \Delta_{2}$ for $n=2 N_{1}+N_{2}+2$ to $2 N_{1}+2 N_{2}+1$.


Figure 1.

The initial values are $x_{1}=0$ and $t_{1}=0$ at $x_{1}=0$. The values of $x_{j}$ and $t_{j}$ at regular points are found by using a numerical integration routine to evaluate Eq. (1). A typical step in MATLAB is $x(j)=x(j-1)+a b s\left(q u a d l\left(F \_x, z(j-1), z(j)\right)\right)$. This scheme is used to evaluate points 2 to $N_{1}$, then the upper turning point is passed by using Eq. (2) to evaluate the $z$ and $t$ values at points $N_{1}+1$ and $N_{1}+2$. The next phase is evaluation of the regular points $N_{1}+3$ to $2 N_{1}+1$. From there, the procedure is replicated, with Eq. (2) used to evaluate points $2 N_{1}+N_{2}+1$ and $2 N_{1}+N_{2}+2$. For points beyond $2 N_{1}+2 N_{2}+1$, the data is replicated according to $x_{n}=x_{2 N_{1}+2 N_{2}+1}+x_{n-2 N_{1}-2 N_{2}-1}$ and $t_{n}=t_{2 N_{1}+2 N_{2}+1}+t_{n-2 N_{1}-2 N_{2}-1}$.

The ray path in Fig. 2 is constructed by plotting $\left(z_{n}, x_{n}\right)$ pairs such that $z$ is displayed as a function of $x$. The first graph shows that the path is moderately curved, with a depth variation of approximately 310 m over a range of more than 4 km . The result for travel time along the ray appears to depend linearly on the range. This property stems from the facts that the grazing angle is small along the entire path and the phase velocity $c(z)$ varies little. Consequently, the horizontal component of the phase velocity which is $c(z) \cos (\chi)$ is nearly constant.


Figure 2.

The path appears to be sinusoidal. Whether this is the true nature of the path is tested by plotting a sine curve relative to the average value of $z$, which is $z_{\mathrm{av}}=$ $[\max (z)+\min (z)] / 2$. The fitted curves are $z=z_{\mathrm{av}}+A \sin \left[2 \pi\left(x-x_{0}\right) / L\right]$ and $t=\max (t) b$, where $A=\left(z_{D}-z_{B}\right) / 2, L / 2$ is the horizontal distance between adjacent locations at which $z=z_{\mathrm{av}}$, and $x_{0}$ is the minimum range at which $z=z_{\mathrm{av}}$ with $d z / d x>0$. This sine curve is indistinguishable from the computed path.

The ray corresponding to $\chi_{\mathrm{I}}=-15.636^{\circ}$ is described by Fig. 3. Launching a ray at an upward or downward initial angle that exceeds this value will generate a ray that reflects from the free surface.


Figure 3.

### 11.2.2 Intensity and Focusing Factor

Thus far, we have not examined how the pressure magnitude is altered as a signal propagates. Although we applied Snell's law for an interface between different fluids to derive equations describing the ray paths, we cannot extend the analogy by applying the reflection and transmission coefficients. The reason for this is that the fluid properties vary continuously. Because $\rho c / \sin \chi$ changes by an infinitesimal amount across a layer of depth $d z$, the reflection coefficient is essentially zero and the transmission coefficient is essentially one. Our need is to determine how much the transmission differs from one, so we shall pursue a different approach. The analysis examines the manner in which power flows along rays in order to obtain an expression for intensity.

A common measurement in the ocean is mapping of the field from a small source. Typically the source radiates axisymmetrically relative to the vertical, which means that the field is independent of the circumferential angle of a cylindrical coordinate system whose axial direction is vertical. This situation is depicted in Fig. 11.10. The ray path in this picture is described by giving the transverse distance $R$ as a function of $z$. The fact that the properties are the same for any vertical plane containing the $z$-axis, combined with the locally planar nature of the geometrical acoustics approximation, requires that $R$ depends on $z$ in the same manner as $x$ depends on $z$ for two-dimensional Cartesian coordinates. In addition, a ray path is altered if the depth $z_{0}$ of the source or the grazing angle at the source are altered. Thus, we may consider the ray path to be defined functionally as $R=R\left(z, z_{0}, \chi_{\mathrm{I}}\right)$.

The analysis of intensity is founded on the concept of a ray tube, which is the region enclosed by a group of rays that leave a source at orientations that are infinitesimally different. A typical tube appears in Fig. 11.10. The sides of this ray tube lie in vertical

Fig. 11.10 A ray tube consisting of all rays emanating from a point at grazing angles between $\chi_{\mathrm{I}}$ and $\chi_{\mathrm{I}}+d \chi_{\mathrm{I}}$, and between azimuthal angles between $\theta$ and $\theta+d \theta$

planes whose circumferential angle differ by $d \theta$. The upper and lower bounding surfaces correspond, respectively, to initial grazing angles $\chi_{\mathrm{I}}$ and $\chi_{\mathrm{I}}+d \chi_{\mathrm{I}}$. The transverse distance to a cross section $d \mathcal{A}$ of the tube is $R$. The cross section is a rectangle whose sides are $R d \theta$ horizontally and $d h$ in the vertical plane.

The significance of a ray tube stems from the fact that the particle velocity, and therefore the intensity, is parallel to a ray. Consequently, there is no power flow across the sides, so the time-averaged power that flows across any cross section of the ray tube must be constant. The particle velocity is parallel to each ray, so the local intensity is perpendicular to the cross section at which it occurs. It follows that the constant power in a ray tube is such that

$$
\begin{equation*}
d \mathcal{P}=\frac{|P|^{2}}{2 \rho_{0} c} d \mathcal{A} \text { is constant along a ray } \tag{11.2.15}
\end{equation*}
$$

An equivalent way of stating this principle is that the time-averaged power that flows across any two cross sections of the same ray tube must be the same. A corollary is that the power radiated by a source into the tube must flow across the cross section at the far end of the tube. Our objective is to describe this principle in terms of the properties of the rays in order to obtain an expression for the intensity.

The cross section of a ray tube in the vicinity of a specified depth is depicted in Fig. 11.11. Both ray paths in this Figure have been terminated at the same depth, so the transverse distance to the end of the upper ray is $R\left(z, z_{0}, \chi_{\mathrm{I}}\right)$, while it is $R\left(z, z_{0}, \chi_{\mathrm{I}}+d \chi_{\mathrm{I}}\right)$ for the lower ray.

Because power is the area integral of intensity, we seek an expression for the cross-sectional area $d \mathcal{A}=(R d \theta)(d h)$. The side $d h$ forms one side of a right triangle whose hypotenuse is the line connecting the points in the bounding rays at depth $z$. The horizontal distance between the ends is solely a consequence of the dependence of $R$ on the initial grazing angle. These angles are $\chi_{\mathrm{I}}$ and $\chi_{\mathrm{I}}+d \chi$ for the rays in the sketch, from which it follows that the horizontal distance between points on the rays is $\left|R\left(z, z_{0}, \chi_{\mathrm{I}}+d \chi\right)-R\left(z, z_{0}, \chi_{\mathrm{I}}\right)\right|$. Therefore, we have

Fig. 11.11 Side view of a ray tube generated by a source at depth $z_{0}$


$$
\begin{equation*}
d h=\left|R\left(z, z_{0,}, \chi_{\mathrm{I}}+d \chi_{\mathrm{I}}\right)-R\left(z, z_{0}, \chi_{\mathrm{I}}\right)\right||\sin \chi|=\left|\frac{\partial R}{\partial \chi_{\mathrm{I}}} d \chi_{\mathrm{I}}\right||\sin \chi| \tag{11.2.16}
\end{equation*}
$$

Thus, the cross-sectional area of the ray tube at the output end is

$$
\begin{equation*}
d \mathcal{A}=(R d \theta) d h=\left|\frac{\partial R}{\partial \chi_{\mathrm{I}}} \sin \chi\right| R d \chi_{\mathrm{I}} d \theta \tag{11.2.17}
\end{equation*}
$$

The normal to $d \mathcal{A}$ is parallel to the rays within the tube, and therefore parallel to the particle velocity on that cross section. Hence, the time-averaged power flowing across $d A$ is

$$
\begin{equation*}
d \mathcal{P}=I_{\mathrm{av}} d \mathcal{A}=I_{\mathrm{av}}\left|\frac{\partial R}{\partial \chi_{\mathrm{I}}} \sin \chi\right| R d \chi_{\mathrm{I}} d \theta \tag{11.2.18}
\end{equation*}
$$

The preceding is the time-averaged power that passes the cross section of the ray tube at depth $z$. This must be the same as the power that is input to the ray tube by the source. To describe this quantity, we observe that heterogeneous effects are significant only over distances that are many wavelengths. Thus, if we surround the source with a small sphere of radius $\varepsilon$, the pressure on this sphere will be the same as it would be if the fluid was homogeneous. If a sphere radiates omnidirectionally, its power $\mathcal{P}$ is distributed evenly over the surface area $4 \pi \varepsilon^{2}$, so the intensity would be a constant value, $\mathcal{P} /\left(4 \pi \varepsilon^{2}\right)$. The radial intensity for a source that is axisymmetric relative to the vertical axis is defined by multiplying this constant reference value by a directivity factor $\mathcal{D}_{I}$. This factor depends on the polar angle, which is the complement of the grazing angle, $\psi=\pi / 2-\chi_{\mathrm{I}}$. Thus, the intensity at radial distance $\varepsilon$ from the source is given by

$$
\begin{equation*}
I_{0}\left(\varepsilon, \chi_{\mathrm{I}}\right)=\frac{\mathcal{P}}{4 \pi \varepsilon^{2}} \mathcal{D}_{I}\left(\pi / 2-\chi_{\mathrm{I}}\right) \tag{11.2.19}
\end{equation*}
$$

The power input $d \mathcal{P}_{0}$ to the ray tube is the product of $I_{0}$ and the area of a patch of the surrounding sphere, whose sides are $\varepsilon d \chi_{\mathrm{I}}$ in the vertical plane and $\left(\varepsilon \cos \chi_{\mathrm{I}}\right) d \theta$ horizontally. Thus, we have

$$
\begin{equation*}
d \mathcal{P}_{0}=I_{0}\left(\varepsilon, \chi_{\mathrm{I}}\right) \varepsilon^{2}\left(\cos \chi_{\mathrm{I}}\right) d \chi_{\mathrm{I}} d \theta=\frac{\mathcal{P}}{4 \pi} \mathcal{D}_{I}\left(\pi / 2-\chi_{\mathrm{I}}\right)\left(\cos \chi_{\mathrm{I}}\right) d \chi_{\mathrm{I}} d \theta \tag{11.2.20}
\end{equation*}
$$

Equating this quantity to the power flowing across an arbitrary cross section in Eq.(11.2.18) yields

$$
\begin{equation*}
I_{\mathrm{av}}=\left(\frac{\mathcal{P}}{4 \pi}\right) \frac{\mathcal{D}_{I}\left(\pi / 2-\chi_{\mathrm{I}}\right) \cos \chi_{\mathrm{I}}}{R\left|\left(\frac{\partial R}{\partial \chi_{\mathrm{I}}}\right) \sin \chi\right|} \tag{11.2.21}
\end{equation*}
$$

The same concept may be employed to examine the intensity when a field is generated by some agent other than a small source. If we know the intensity $I_{1}$ along a cross section at one location, and we wish to determine the value at another cross section, equality of the power flows requires that

$$
\begin{equation*}
d \mathcal{P}=\left(I_{\mathrm{av}}\right)_{2} d A_{2}=\left(I_{\mathrm{av}}\right)_{1} d A_{1} \tag{11.2.22}
\end{equation*}
$$

When we use Eq. (11.2.18) to describe each term, we obtain

$$
\begin{equation*}
\left(I_{\mathrm{av}}\right)_{2}=\left(I_{\mathrm{av}}\right)_{1} \frac{R_{1}\left|\left(\frac{\partial R}{\partial \chi_{\mathrm{I}}}\right)_{1} \sin \chi_{1}\right|}{R_{2}\left|\left(\frac{\partial R}{\partial \chi_{\mathrm{I}}}\right)_{2} \sin \chi_{2}\right|} \tag{11.2.23}
\end{equation*}
$$

If we know the ray path, we can evaluate all terms in this expression other than the value at the input end. Thus, we can use this relation to monitor the intensity as the signal propagates along the ray path provided that we have a starting value for the intensity.

The focusing factor is used to compare the intensity to the hypothetical value that would be obtained at a specified location if the same source radiated into a homogeneous fluid. The radial distance would be $r=\left(R^{2}+x^{2}\right)^{1 / 2}$. Equation(11.2.19) describes the intensity for a homogeneous fluid, so we may find the reference intensity by replacing $\varepsilon$ in that relation by $r$. Thus, the focusing factor is

$$
\begin{equation*}
\frac{I_{\mathrm{av}}}{I_{0}\left(r, \chi_{\mathrm{I}}\right)}=\frac{r^{2} \cos \chi_{\mathrm{I}}}{R\left|\frac{\partial R}{\partial \chi_{\mathrm{I}}} \sin \chi\right|} \tag{11.2.24}
\end{equation*}
$$

The condition where $\partial R / \partial \chi_{\mathrm{I}}=0$ marks a location at which the cross-sectional area of a ray tube narrows to zero. The actual pressure must be finite, so this condition marks a condition where ray theory requires some adjustment. Depending on the specific nature of the signal, it might be that the singularity may be corrected by using a higher order approximation within the context of a linear acoustic theory, or it might be appropriate to incorporate a nonlinear correction. In any event the pressure will be greatly enhanced in the vicinity of a location where $\partial R / \partial \chi_{\mathrm{I}}=0$.

Extreme narrowing of ray tubes is manifested in either of two ways. One is a focus, like that encountered in an optical lens. A focus is marked by convergence


Fig. 11.12 Two situations in which geometric acoustics predicts a singularity of intensity: a a focus, $\mathbf{b}$ a caustic
of a group of rays on a single point. According to geometrical acoustics, the rays emerge from the focus continuously, unless the focus is at an interface between media. Figure 11.12a displays a focus at which a cone of rays intersect. The quantity $\partial R / \partial \chi_{\mathrm{I}}$ is zero because there is no change in the transverse distance $R$ for rays associated with initial grazing angles that differ by an infinitesimal amount. This is different from the situation where two rays associated with finitely different $\chi_{\mathrm{I}}$ values intersect. In that case, the pressure at the intersection is the (finite) sum of the pressure associated with each ray.

Figure 11.12b illustrates a different singularity. In it, there is a curve in two dimensions, or a surface in three dimensions, on which rays for infinitesimally different $\chi_{\mathrm{I}}$ arrive tangentially. We previously encountered this phenomenon in Example 11.1. The line or surface that is the locus of points on which $R$ does not change when $\chi_{\mathrm{I}}$ is changed infinitesimally is a caustic. When the fluid properties and the starting location of a ray are specified, the transverse distance $R$ to a point on a ray path depends only on the depth and the initial grazing angle, that is, $R=f\left(z, \chi_{\mathrm{I}}\right)$. Hence, the occurrence of a caustic in the axisymmetric field in a vertically stratified fluid is marked by

$$
\begin{equation*}
R=f\left(z, \chi_{\mathrm{I}}\right) \text { and } \frac{\partial f}{\partial \chi_{\mathrm{I}}}=0 \tag{11.2.25}
\end{equation*}
$$

In principle, we could identify this locus by solving the second equation for $z$ over a range of $\chi_{\mathrm{I}}$ values, then using the first equation to find the corresponding $R$ values. In practice, the ray path usually is determined numerically, so the function $f$ will not be available. A caustic is marked visually by a greatly reduced spacing between adjacent rays, with all arriving rays being on one side of the caustic. The absence of rays on the other side of the caustic is an artifice of the geometrical acoustics approximation. The actual behavior is that the field evanesces very rapidly in the direction normal to the caustic. Nevertheless, geometrical acoustics is quite useful even here because it informs us that there is a region where the pressure is greatly enhanced. Furthermore, Eq. (11.2.23) may be used to evaluate the intensity beyond the caustic. (This extension is a topic in Sect.11.3.2.)


Fig. 11.13 Ray paths for a point source in a channel modeled by Munk in the Journal of the Acoustical Society of America, vol. 55, pp. 220-226 (1974). The dependence of sound speed on depth is $c=c_{1}\left(1+\varepsilon\left(\xi+e^{-\zeta}-1\right)\right)$ where $\xi=2\left(z-z_{1}\right) / z_{1}$. The coefficients are $c_{1}=1492 \mathrm{~m} / \mathrm{s}$, $\varepsilon=0.0074$, and $z_{1}=1300 \mathrm{~m}$

A classic demonstration of the occurrence of caustics may be found in a paper by Munk. ${ }^{8}$ The sound speed profile described in Fig. 11.13 was suggested there to be a good prototype for studies of phenomena in the ocean at moderate latitudes. The sound speed profile has a minimum at $z_{1}=1300 \mathrm{~m}$. An omnidirectional source is located at this depth, so that $z_{\mathrm{I}}=z_{1}$. Setting $d c / d z=0$ at $z_{\mathrm{I}}$ results in the ray paths having an inflection point at the depth of the source. However, the sound speed profile is not symmetric relative to this depth. Consequently, the appearance of a ray path below $z_{\mathrm{I}}$ does not mirror that appearance above.

The rays appearing in Fig. 11.13 correspond to initial grazing angles ranging from $14.40^{\circ}$ upward to $14.40^{\circ}$ downward. This is the range of angles for which rays are not reflected from the surface. The rays described here extend to a greater range than those computed by Munk, whose results were not shown in their entirety beyond a range of 25 km .

The most evident feature of the ray paths is that there are regions where there are none. These are shadow zones. Usually, one thinks of a shadow as the blocking effect of a solid object. Here, a caustic obstructs the sound field. Regions of enhanced intensity correspond to decreased spacing between adjacent rays. They appear as darkened regions in the figure. The source is at $R=0, z=1300 \mathrm{~m}$, so the enhancement near that location is not surprising. The next enhanced region occurs in the range from 10 to 20 km at depths between $400 \mathrm{~m}-1.3 \mathrm{~km}$. This appears to be a caustic. Other regions of interest are close to the source depth at ranges in the vicinity of 50 and 70 km . These seem to be somewhat like a focus, but they also seems to resemble a caustic. Figure 11.14 zooms in on the regions where rays are concentrated.

[^52]

Fig. 11.14 Zoomed views of regions of enhanced intensity in the sound channel studied by Munk in the Journal of the Acoustical Society of America, vol. 55, pp. 220-226 (1974)

The caustic in the left figure is a boundary beyond which the pressure is zero (or very small if the geometrical acoustics approximation is corrected). The low point of this feature occurs at the depth of the source. The concentration of rays in the vicinity of $R=47 \mathrm{~km}$ and $R=70 \mathrm{~km}$ in the right figure also occurs in the vicinity of the source depth $z_{\mathrm{I}}$. The zoomed views show that these concentrations are not foci, because only rays at nearly equal grazing angles intersect. This is a fundamental aspect of caustics. Each has a distinct arrete,which is the minimum distance for the caustic. The rays become more diffuse at the other end of each caustic. Additional computations would show that caustics like these will occur even if $z_{\mathrm{I}}$ is not the same as depth $z_{1}$ for minimum $c(z)$. However, in that case no caustics would be observed at $z_{1}$ or $z_{1}$.

EXAMPLE 11.3 A radially symmetric point source is situated at $z_{\mathrm{I}}=150$ m . The sound speed dependence on depth is as described in Example 11.2. Identify the regions where the pressure is enhanced. Then explain how this data may be used to evaluate the focusing factor at a designated location.

## Significance

In addition to showing how to map rays, some interesting phenomena, including the role of reflections, will be explored.

## Solution

We begin with an evaluation of the ray paths that remain below the surface. For this, we follow the procedure developed in the previous example. Stated briefly, the
initial grazing angle is set in a range whose limits are the values extracted from Eq. (11.2.8), specifically $\chi_{\mathrm{I}}=\cos ^{-1}\left(c_{\mathrm{I}} / c(z=0)\right)= \pm 15.363^{\circ}$. The turning point depths $\min (z)$ and $\max (z)$ for each angle are computed, and the depth is discretized into intervals of $\Delta_{1}$ for $\min (z) \leq z \leq z_{\mathrm{I}}$ and $\Delta_{2}$ for $z_{\mathrm{I}} \leq z \leq \max (z)$. Starting from the source location, the transverse value of $R$ at each depth up to the first turning point is computed by applying the first of Eq. (1) in Example 11.2. The turning point and the next are evaluated by applying Eq.(2). Then, Eq.(1) is used to evaluate points up to the second turning point, followed by Eq. (2) for that point and the point beyond. The final part of evaluation uses Eq. (1) to return to $z_{\mathrm{I}}$, after which the path is periodically replicated. Figure 1 is the result of this computation.


Figure 1. Rays radiated by the source that do not reflect from the surface.

It can be seen that all rays have the same periodicity horizontally, and that they all intersect at a common point. These intersections are foci. Both properties are a consequence of $c(z)$ being an even function relative to the depth $z_{\min }$ at which it is a minimum. Another consequence of the symmetric nature of the sound speed profile is that the depth of foci are either $z_{\mathrm{I}}$ or $\max (z)-z_{\mathrm{I}}$.

In addition to the occurrence of foci, there also appears to be large shadow zones. However, Fig. 1 does not give a complete picture, because it does not account for rays that reflect from the free surface. Such a condition is marked by one of the turning points being above the surface, that is, $z_{0}<0$. Thus, reflected rays in the present case correspond to $15.363^{\circ}<\chi_{\mathrm{I}}<90^{\circ}$ and $-90^{\circ}<\chi_{\mathrm{I}}<-15.363^{\circ}$. Rays that reflect are computed by using Eq. (1) in Example 11.2 to evaluate the values of $R$ and $x$ at the reflection point. If we let $J$ equal to the index of the reflection point, then the properties for the next point are set to place its properties symmetrically, so that $R_{J+1}=R_{J}+\left(R_{j}-R_{J-1}\right)$.

The result of including rays that reflect is Fig. 2. It appears that only one shadow zone remains. (Even that zone would disappear if there were a bottom at a finite depth, in which case there would also be reflections at that boundary.) The ray paths are far apart in the regions that were quiet in the previous figure. This tells us that the intensity is low in those regions. There appears to be a caustic that touches the surface at a range of 2.5 km formed from a group of rays for which $\chi_{\mathrm{I}}>0$, and another at a range of 2.8 km formed from a group for which $\chi_{\mathrm{I}}<0$.


Figure 2. All rays radiated by the point source. Blue paths are rays that do not relect from the surface, red rays radiate upward from the source $\left(\chi_{\mathrm{I}}<0\right)$, and green rays radiate downward $\left(\chi_{\mathrm{I}}>0\right)$.

To examine these features more closely, Fig. 3 displays only paths of rays that reflect from the surface. The details remain masked in the overview, but they can be seen clearly in the zoomed views. The broad darkened region in the vicinity of the source is due to nearly vertical rays that reflect from the surface. They return to the depth of the source at a comparatively small horizontal distance, so they enhance the already large pressure in the nearfield of a spherical source. In the zoomed view of the horizontal range from 2 to 3.2 km , we see that there are indeed two caustics formed from rays that are incident at the surface. We also see that reflection of the rays associated with each caustic seems to result in each forming another caustic.


Figure 3. Rays that reflect from the free surface when the sound speed depends parabolically on depth.

Graphs such as these give a qualitative picture, but the data also may be used to calculate the focusing factor. The quantity in Eq. (11.2.24) that is not readily available is $\partial R / \partial \chi_{\mathrm{I}}$. Because $z$ is held constant in this derivative, we could estimate it with a finite difference if we had values of $R$ corresponding to a discretization of $z$ that is common to the ray paths. The problematic aspect is that our algorithm for evaluating ray paths leads to the type of data depicted in Fig.4. The two ray paths correspond to slightly different initial grazing angles $\chi_{1}$ and $\chi_{2}$. Points marked $a$ and $b$ on each curve are the result of computations, so the $R$ and $z$ coordinates of each point are known. The algorithm that we use sets an equal depth increment for the computation of a ray, but that increment is not constant from ray to ray. Thus, $z_{a}\left(\chi_{2}\right) \neq z_{a}\left(\chi_{1}\right)$ and $z_{b}\left(\chi_{2}\right) \neq z_{b}\left(\chi_{1}\right)$.


Figure 4.

We wish to determine the intensity at the indicated depth $z$ midway between these ray paths. To that end we first evaluate the transverse distance to the point on each ray path at depth $z$. This is done by interpolating between the coordinates of the computed point. A linear interpolation gives

$$
R\left(z, \chi_{j}\right)=\left.\left(\frac{z_{b}-z}{z_{b}-z_{a}} R_{a}+\frac{z-z_{a}}{z_{b}-z_{a}} R_{b}\right)\right|_{\chi_{j}}
$$

Then a first-order finite difference approximation gives

$$
\left.\frac{\partial R}{\partial \chi_{\mathrm{I}}}\right|_{z,\left(R_{a}+R_{b}\right) / 2}=\frac{R\left(z, \chi_{2}\right)-R\left(z, \chi_{1}\right)}{\chi_{2}-\chi_{2}}
$$

Application of this procedure to create an intensity map for the entire field would require effort. One approach would establish a spatial grid at which the algorithm would be implemented. A search algorithm would identify the two rays that are closest to a selected grid point, and the data points on each ray that are directly above and below the selected point. Section 11.3.2 will develop a formulation that is readily programed and more accurate.

### 11.3 Arbitrary Heterogeneous Fluids

The model of a vertically stratified fluid is useful for environmental acoustic studies. However, it might be that the sound speed depends on the horizontal range, as well as depth, especially at very long ranges. The analytical approach in the previous section has the attractive feature of being pictorial, and therefore readily understood. The method that follows is mathematical and more abstract, but it is general.

### 11.3.1 Ray Tracing Equations

The formulation begins with the general equation for the zero-phase wavefront, which was stated previously in Eq.(11.1.4),

$$
\begin{equation*}
\tau(\bar{x}) \equiv \frac{\bar{x} \cdot \bar{e}(\bar{x})}{c(\bar{x})}=t \tag{11.3.1}
\end{equation*}
$$

The corresponding pressure is given by Eq.(11.1.1) for a harmonic signal or Eq.(11.1.8) for an arbitrary time signature. In either case, a constant value of $\tau$ locates the wavefront on which the phase is zero when $t$ is that value of $\tau$.

The geometrical acoustics approximation of propagation was depicted in Fig. 11.1, where a specific signal is observed at position $\bar{x}=\bar{\xi}$ at time $t$. At a slightly later instant $t+d t$, it is situated on the same ray at $\bar{x}=\bar{\xi}+d \bar{\xi}$. Because $\bar{\xi}$ marks the position of a specific phase, it must be that $d \bar{\xi} / d t$ is the phase velocity. A different view is that the local sound speed is $c(\bar{\xi})$, and the propagation is in the direction $\bar{e}(\bar{\xi})$ tangent to the ray path. Both views are correct, so it must be that

$$
\begin{equation*}
\frac{d \bar{\xi}}{d t}=c(\bar{\xi}) \bar{e}(\bar{\xi}) \tag{11.3.2}
\end{equation*}
$$

Another view of the propagation is that the phase moves along a ray such that it remains on the same wavefront. The wavefront at time $t$ is described by Eq. (11.3.1), and its shape at time $t+d t$ is described by $\tau(\bar{\xi}+d \bar{\xi})=t+d t$. A Taylor series expansion this expression gives

$$
\begin{equation*}
\tau(\bar{\xi})+\nabla \tau \cdot d \bar{\xi}=t+d t \tag{11.3.3}
\end{equation*}
$$

The fact that $\tau(\bar{\xi})=t$ reduces this relation to

$$
\begin{equation*}
\nabla \tau \cdot \frac{d \bar{\xi}}{d t}=1 \tag{11.3.4}
\end{equation*}
$$

A fundamental property of the gradient of a function is that it is normal to the surface on which the function is constant. Along a wavefront $\tau$ is constant, and $\bar{e}(\bar{\xi})$ is the
normal, so we have $\nabla \tau=|\nabla \tau| \bar{e}(\bar{\xi})$. Substitution of this description and Eq. (11.3.2) into the preceding leads to $|\nabla \tau|=1 / c(\bar{\xi})$. An equivalent description of this property is

$$
\begin{equation*}
\nabla \tau \cdot \nabla \tau=\frac{1}{c(\bar{\xi})^{2}} \tag{11.3.5}
\end{equation*}
$$

This is the eikonal equation. Because the magnitude of $\nabla \tau$ is inversely proportional to the local sound speed, and it is in the sense of the propagation, it is called the wave slowness. We shall denote this quantity as $\bar{s}(\bar{\xi})$, so we have

$$
\begin{equation*}
\bar{s}(\bar{\xi}) \equiv \nabla \tau=\frac{\bar{e}(\bar{\xi})}{c(\bar{\xi})} \tag{11.3.6}
\end{equation*}
$$

Manipulation of the relations developed thus far will lead to a set of first-order differential equations in which the dependent variables are $\bar{\xi}$ and $\bar{s}$. The derivation begins by evaluating the rate at which $\bar{s}$ changes. One way in which that derivative may be found is by the chain rule based on considering $\bar{s}$ to be a function of position along a ray, which leads to

$$
\begin{equation*}
\frac{d \bar{s}}{d t}=\nabla \bar{s} \cdot \frac{d \bar{\xi}}{d t}=\nabla \bar{s} \cdot(c \bar{e}) \tag{11.3.7}
\end{equation*}
$$

(The quantity $\nabla \bar{s}$ is the gradient of a vector, which we usually do not encounter. In Cartesian coordinates, each component of $\bar{s}$ may depend on $x, y$, and $z$, and the unit vector parallel to each axis is constant. Correspondingly, we would find that $\nabla \bar{s}=\nabla s_{x} \bar{e}_{x}+\nabla s_{y} \bar{e}_{y}+\nabla s_{z} \bar{e}_{z}$.) We use the definition of $\bar{s}$, Eq.(11.3.6), which indicates that $\bar{e}=c \bar{s}$, to eliminate the unit vector,

$$
\begin{equation*}
\frac{d \bar{s}}{d t}=c^{2} \nabla \bar{s} \cdot \bar{s} \equiv \frac{1}{2} c^{2} \nabla(\bar{s} \cdot \bar{s}) \tag{11.3.8}
\end{equation*}
$$

The eikonal equation gives $\bar{s} \cdot \bar{s}=1 / c^{2}$. The gradient of this quantity is $\nabla(\bar{s} \cdot \bar{s})=$ $\left(-2 / c^{3}\right) \nabla c$. Substitution of this relation converts the preceding equation to

$$
\begin{equation*}
\frac{d \bar{s}}{d t}=-\frac{\nabla c}{c(\bar{\xi})} \tag{11.3.9}
\end{equation*}
$$

In addition, substituting $\bar{e}=c \bar{s}$ into Eq. (11.3.2), which describes the phase velocity, leads to

$$
\begin{equation*}
\frac{d \bar{\xi}}{d t}=c(\bar{\xi})^{2} \bar{s}(\bar{\xi}) \tag{11.3.10}
\end{equation*}
$$

Equations (11.3.9) and (11.3.10) are coupled ordinary differential equations that govern $\bar{\xi}$ and $\bar{s}$ as functions of time. They are known as the ray tracing equations. Except for a few special cases, their solution requires computational methods. The numerical algorithms typically implemented in software solve first-order equations that have been placed into a matrix form,

$$
\begin{equation*}
\frac{d}{d t}\{X\}=\{F\} \tag{11.3.11}
\end{equation*}
$$

The ray tracing equations are expressed in this form by placing the components of $\bar{s}$ and $\bar{\xi}$ in the column vector $\{X\}$. The components of the right side of each equation are placed in the same sequence in $\{F\}$. The result is a set of six differential equations,

$$
\begin{align*}
& \{X\}=\left[\begin{array}{llllll}
s_{x} & s_{y} & s_{z} & \xi_{x} & \xi_{y} & \xi_{z}
\end{array}\right]^{\mathrm{T}} \\
& \{F\}=\left[-\frac{1}{c}\left[\begin{array}{lll}
\frac{\partial c}{\partial \xi_{x}} & \frac{\partial c}{\partial \xi_{y}} & \frac{\partial c}{\partial \xi_{z}}
\end{array}\right] \quad c^{2}\left[\begin{array}{lll}
s_{x} & s_{y} & s_{z}
\end{array}\right]\right]^{\mathrm{T}} \tag{11.3.12}
\end{align*}
$$

The sound speed is a function of position. Hence, it is implicit that embedded in the computation of $\{F\}$ are routines that evaluate $c$ and the components of its gradient as functions of the position coordinates $\xi_{x}, \xi_{y}$, and $\xi_{z}$.

To start the numerical solution, the initial conditions must be stated. The initial instant may be set as $t=0$ because time does not appear explicitly in the ray tracing equations. The analysis of rays in a vertically stratified fluid began by setting the starting position and the grazing angle. The analogous information here is a specification of where a ray is situated, $\bar{\xi}=\bar{\xi}_{0}$, and the direction in which it is aimed, $\bar{e}=\bar{e}_{0}$, at $t=0$. The corresponding initial wave slowness is $\bar{s}_{0}=\bar{e}_{0} / c\left(\bar{\xi}_{0}\right)$. The result of the numerical solution is a set of $\bar{\xi}$ and $\bar{s}$ values at a set of $t$ value for a single ray. Depending on how the signal is generated, we might find families of rays by changing the starting point or the initial tangent direction.

It generally is a good idea to have an external check to monitor the error when differential equations are solved numerically. For example, if energy is conserved in a dynamic system, the mechanical energy contained in the response at each instant can be verified to be constant. A conservation-like principle is the eikonal equation, $\bar{s} \cdot \bar{s}=1 / c^{2}$. Solution of the differential equations gives the values of $\bar{\xi}$ and $\bar{s}$ at some instant $t$. Knowledge of these values makes it possible to compute $c(\bar{\xi})^{2} \bar{s} \cdot \bar{s}-1$. The degree to which the computed value differs from zero is a fractional measure of the error.

One situation where the ray tracing equations may be solved analytically is that of a homogeneous fluid, for which $\nabla c \equiv 0$. Equation (11.3.9) states that $s$ is constant in this case, so it never changes from its initial value, $\bar{s}_{0}$. In turn, this means that the right side of Eq.(11.3.10) is constant at the initial value $\bar{s}_{0}$. The solution of that equation is

$$
\begin{equation*}
\bar{\xi}=\bar{\xi}_{0}+c^{2} \bar{s}_{0} t=\bar{\xi}_{0}+\bar{e}\left(\bar{\xi}_{0}\right) c t \tag{11.3.13}
\end{equation*}
$$

This is the equation of a straight line. Thus, the ray tracing equations for a homogeneous fluid yield rays that are straight, which is a property we already knew. Furthermore, the signal at instant $t$ is at distance $c t$ from the starting point, as it must for plane wave propagation.

EXAMPLE 11.4 Consider a fluid whose sound speed depends linearly on the depth. Start with an ansatz for a ray path that is the general equation for a circle. Show that the ray tracing equations are satisfied by this representation if the center of the circle is situated at the extrapolated elevation where the value of $c$ is zero.

## Significance

The process of working backward to fit the trial solution to the ray tracing equations will clarify what the forward solution process actually entails. It also serves to tie the graphical approach for vertically stratified media to the mathematical formulation for arbitrary heterogeneity.

## Solution

The mathematical operations are somewhat simplified by defining $z=0$ to be the depth at which $c=0$. (This location might be outside the region in which the fluid resides.) This definition of $z$ leads to $c=\beta z$, where $\beta$ is an arbitrary constant. There are several ways in which a circular path may be represented mathematically. We will use a parametric form in which the grazing angle $\chi$ is the parameter. Figure 1 shows that $\chi$ also is the angle of the radial line to a point on the ray. The coordinates of the center are $\left(x_{0}, z_{0}\right)$. Neither these coordinates nor the radius $\mathcal{R}$ is presumed to be known.


Figure 1.

The diagram shows that the coordinates of a point on the circular arc are

$$
\begin{equation*}
\xi_{x}=x_{0}+\mathcal{R} \sin \succ, \quad \xi_{z}=z_{0}-\mathcal{R} \cos \succ \tag{1}
\end{equation*}
$$

Thus, the sound speed at $\bar{x}=\bar{\xi}$ is

$$
\begin{equation*}
c(\bar{\xi})=\beta\left(z_{0}-\mathcal{R} \cos \succ\right) \tag{2}
\end{equation*}
$$

The slowness vector is tangent to the ray, and its magnitude is $1 / c$, so we find from the sketch that

$$
\begin{equation*}
\bar{s}=\frac{1}{c(\bar{\xi})}\left(\bar{e}_{x} \cos \succ+\bar{e}_{z} \sin \chi\right) \tag{3}
\end{equation*}
$$

The last aspect of the trial solution is the introduction of time. A signal travels along a ray at the local speed of sound. This is the rate at which point $\vec{\xi}$ moves along the arc. The path is circular, so the corresponding angular speed is

$$
\begin{equation*}
\frac{d \chi}{d t}=\frac{c(\bar{\xi})}{\mathcal{R}} \tag{4}
\end{equation*}
$$

Our approach is to substitute these representations into the ray tracing equations. Because all quantities have been defined in terms of $\chi$, derivatives with respect to time may be evaluated by the chain rule. We begin with Eq.(11.3.10), for which we have

$$
\frac{d \bar{\xi}}{d t}=\left(\frac{d \xi_{x}}{d \chi} \bar{e}_{x}+\frac{d \xi_{z}}{d \chi} \bar{e}_{z}\right) \frac{d \chi}{d t}=\mathcal{R}\left(\bar{e}_{x} \cos \succ+\bar{e}_{z} \sin \succ\right)\left(\frac{c}{\mathcal{R}}\right) \equiv c^{2} \bar{s}
$$

This shows that Eq. (1) identically satisfies Eq. (11.3.10) independently of the radius and coordinates of the center. In other words, any circular path is consistent with that equation.

Consideration of Eq. (11.3.9) proceeds similarly. An expression for $d s / d t$ may be found by differentiating Eq. (3). The key observation is that Eq. (2) gives $c$ as a function of $\chi$, which means that Eq. (3) actually is a description of $s$ as a function of $\chi$. This calls for the chain rule for differentiation, which leads to

$$
\begin{aligned}
\frac{d \bar{s}}{d t}= & \left(\frac{d \chi}{d t}\right)\left(\frac{d \bar{s}}{d \chi}\right)=\frac{c}{\mathcal{R}}\left[\frac{d}{d \chi}\left(\frac{1}{c}\right)\left(\bar{e}_{x} \cos \chi+\bar{e}_{z} \sin \chi\right)\right. \\
& \left.+\frac{1}{c} \frac{d}{d \chi}\left(\bar{e}_{x} \cos \chi+\bar{e}_{z} \sin \chi\right)\right] \\
= & \left(\frac{c}{\mathcal{R}}\right)\left[-\frac{1}{c^{2}} \frac{d c}{d \chi}\left(\bar{e}_{x} \cos \chi+\bar{e}_{z} \sin \chi\right)+\frac{1}{c}\left(-\bar{e}_{x} \sin \chi+\bar{e}_{z} \cos \chi\right)\right] \\
= & \frac{1}{\mathcal{R}}\left[-\left(\frac{\beta \mathcal{R} \sin \chi}{\beta z_{0}-\mathcal{R} \cos \chi}\right)\left(\bar{e}_{x} \cos \chi+\bar{e}_{z} \sin \chi\right)+\left(-\bar{e}_{x} \sin \chi+\bar{e}_{z} \cos \chi\right)\right]
\end{aligned}
$$

Collection of like components reduces this expression to

$$
\begin{equation*}
\frac{d \bar{s}}{d t}=\frac{\beta}{\beta\left(z_{0}-\mathcal{R} \cos \chi\right) \mathcal{R}}\left[z_{0}\left(-\bar{e}_{x} \sin \chi+\bar{e}_{z} \cos \chi\right)-\mathcal{R} \bar{e}_{z}\right] \tag{5}
\end{equation*}
$$

This expression constitutes the left side of Eq.(11.3.9). The right side is found by using the original description $c=\beta z$ to evaluate the gradient. Thus,

$$
\begin{equation*}
-\frac{\nabla c}{c}=-\left.\frac{1}{c}\left(\frac{\partial c}{\partial z} \bar{e}_{z}\right)\right|_{z=\xi_{x}}=-\frac{1}{\beta\left(z_{0}-\mathcal{R} \cos \chi\right)}\left(\beta \bar{e}_{z}\right) \tag{6}
\end{equation*}
$$

When we equate Eqs. (5) and (6), we see that the $\bar{e}_{x}$ and $\bar{e}_{z}$ components match only if $z_{0}=0$. Because $z_{0}$ is the depth of the center of curvature, the finding that $z_{0}=0$ places the center of the circular ray path at the depth where $c$ vanishes. As is true for the first ray tracing equation, Eq. (11.3.9) is satisfied for any $\mathcal{R}$. This quantity and the actual location of the center are set by the initial position and angle at which the ray is launched. The circular nature of the ray path in a linearly varying fluid was identified previously. Solution of the ray tracing equations gives us a concrete illustration of what the ray tracing equations represent.

### 11.3.2 Amplitude Dependence

## Transport Equation

Suppose we know the waveform of a signal at the starting location $\bar{\xi}_{0}$ of a ray. We wish to know the waveform at a downstream point $\bar{\xi}_{A}$ on that ray. Part of such a determination requires that we find the travel time for the signal. The discussion of Eq. (11.1.4) informs us that we can extract this property directly from the solution of the ray tracing equations. The solution algorithm sets the time base such that $t=0$ for the starting point $\bar{\xi}_{0}$. The wavefronts that are computed correspond to zero phase. This means that the signal that departed from $\bar{\xi}_{0}$ at $t=0$ will be observed at the computed location $\bar{\xi}_{A}$ at time $t_{A}$. According to the aforementioned discussion, a signal that departs from $\bar{\xi}_{0}$ at some instant $t^{\prime}>0$ will arrive at $\bar{\xi}_{A}$ at time $t_{A}+t^{\prime}$. Consequently, if we retain the time value associated with each $\bar{\xi}$ value obtained in a numerical solution of the ray tracing equations, then we may use that time to determine the time delay of the waveform function at each location. A different route leading to the same conclusion comes from the observation that the independent variable for the ray tracing equations is $t$, but it does not appear explicitly anywhere in those equations other than as the rate variable for derivatives. Thus, shifting the time at $\bar{\xi}_{0}$ from $t=0$ to $t=t^{\prime}$ merely increments the time value at any point $\bar{\xi}_{A}$ from $t_{A}$ to $t_{A}+t^{\prime}$.

Knowledge of the time delay allows us to determine when a specific phase arrives at any location along a computed ray. However, this knowledge sheds no light on the amplitude dependence of the signal. Here, we will develop a procedure by which this dependence may be identified if the signal is harmonic. The basic notion is to seek conditions that must be satisfied in order that the geometrical acoustics description of pressure be a solution of the position-dependent wave equation. For this analysis, we use the field equation that accounts for density variation, which is Eq. (11.1.21). Doing so yields a result that is valid for gases, as well as liquids.

We begin with the basic ansatz in Eq. (11.1.1). In term of the wavefront function $\tau(\bar{x})$, it is

$$
\begin{equation*}
p(x, t)=P(\bar{x}) \operatorname{Re}\left(e^{i \omega(t-\tau(\bar{x}))}\right) \tag{11.3.14}
\end{equation*}
$$

From this, we find the pressure gradient to be

$$
\begin{equation*}
\nabla p=\operatorname{Re}\left((\nabla P-i \omega P \nabla \tau) e^{i \omega(t-\tau(\bar{x}))}\right) \tag{11.3.15}
\end{equation*}
$$

The divergence of this expression is

$$
\nabla^{2} p=\operatorname{Re}\left(\left(\nabla^{2} P-2 i \omega \nabla P \cdot \nabla \tau-i \omega P \nabla^{2} \tau-\omega^{2} P \nabla \tau \cdot \nabla \tau\right) e^{i \omega(t-\tau(\bar{x}))}\right)
$$

The dependence of $p$ on $t$ is harmonic. Thus, substitution of the preceding into Eq. (11.1.21) leads to

$$
\begin{align*}
\nabla^{2} p-\frac{1}{\rho_{0}} \nabla \rho_{0} \cdot \nabla p-\frac{1}{c^{2}} \frac{\partial^{2} p}{\partial t^{2}}= & \operatorname{Re}\left\{\left[\nabla^{2} P-2 i \omega \nabla P \cdot \nabla \tau-i \omega P \nabla^{2} \tau\right.\right. \\
& -\omega^{2} P \nabla \tau \cdot \nabla \tau-\frac{1}{\rho_{0}} \nabla \rho_{0} \cdot(\nabla P-i \omega P \nabla \tau) \\
& \left.\left.+\frac{\omega}{c^{2}} P\right] e^{i \omega(t-\tau(\bar{x}))}\right\}=0 \tag{11.3.16}
\end{align*}
$$

This equation must be satisfied for any combination of $\bar{x}$ and $t$ values, so the complex exponential may be canceled. The fact that both $P$ and $\tau$ are real assists decomposition of the factor of this exponential into its real and imaginary parts. The real part is

$$
\begin{equation*}
\nabla^{2} P+\frac{1}{\rho_{0}} \nabla \rho_{0} \cdot \nabla P-\omega^{2}\left(\nabla \tau \cdot \nabla \tau-\frac{1}{c^{2}}\right) P=0 \tag{11.3.17}
\end{equation*}
$$

Geometrical acoustics considers $\omega$ to be very large, which means that terms that are proportional to $\omega^{2}$ are dominant. This attribute makes it permissible to drop the terms that not contain $\omega$. What remains is the factor in parentheses multiplying $P$. The relation must be true for any $P$. Therefore, the term inside the parentheses must be zero. The result is a restatement of the eikonal equation, so it provides no new information.

The imaginary part of Eq. (11.3.16) is the one that is useful. All of these terms are proportional to $\omega$, so none may be dropped as being insignificant relative to others. Factoring out $-\omega$ converts the imaginary part to the transport equation,

$$
\begin{equation*}
2 \nabla P \cdot \nabla \tau+P \nabla^{2} \tau-\frac{1}{\rho_{0}} P \nabla \rho_{0} \cdot \nabla \tau=0 \tag{11.3.18}
\end{equation*}
$$

This expression features three terms, each of which is a derivative of one variable in a triple product. The term $-\left(\nabla \rho_{0}\right) / \rho_{0}$ is identically $\rho_{0} \nabla\left(1 / \rho_{0}\right)$ and the two factor for the first term suggests that the underlying term is $P^{2}$. Consequently, let us multiply
the transport equation by $P / \rho_{0}$. Doing so gives

$$
\begin{equation*}
2 \frac{P}{\rho_{0}} \nabla P \cdot \nabla \tau+\frac{P^{2}}{\rho_{0}} \nabla^{2} \tau+P^{2} \nabla\left(\frac{1}{\rho_{0}}\right) \cdot \nabla \tau \equiv \nabla \cdot\left(\frac{P^{2}}{\rho_{0}} \nabla \tau\right)=0 \tag{11.3.19}
\end{equation*}
$$

This relation is further simplified by the definition that $\nabla \tau \equiv s \equiv \bar{e} / c$. Substitution of this relation reduces the transport equation to

$$
\begin{equation*}
\nabla \cdot\left(\frac{P^{2}}{\rho_{0} c} \bar{e}\right)=0 \tag{11.3.20}
\end{equation*}
$$

The particle velocity in the geometrical acoustic approximation fits the relation for a planar wave, according to which $\bar{V}=\bar{e} P /\left(\rho_{0} c\right)$. Therefore, the quantity inside the parentheses is twice the time-averaged intensity tangent to the ray. If we already have solved the ray tracing equations, then the intensity is the sole unknown in this version of the transport equation. In principle, we could employ a numerical differential equation solver to determine how the intensity varies along the ray path. However, there is a simpler way of extracting $P$ along a ray. It does not require solution of a differential equation.

## Relation to Ray Tube Area

The occurrence of intensity as the variable governed by the transport equation suggests that this relation is fundamentally like the result for intensity in a depthdependent heterogeneity in Sect. 11.2.2. To examine this observation, we consider a ray tube, such as the one in Fig. 11.15. The starting point for the ray tube is $\bar{x}=\bar{\xi}_{0}$, and the cross-sectional area at any location $\bar{\xi}$ is the infinitesimal value $d \mathcal{A}(\bar{\xi})$.

Fig. 11.15 Ray tube in an arbitrary heterogeneous fluid. The edges are rays whose initial tangent corresponds to four sets of direction angles that differ by differential amounts


Green's (divergence) theorem is applied to the transport equation, with the domain $\mathcal{V}$ being the region contained within the ray tube. The intensity is in the direction $\bar{e}$ tangent to the rays, and the normal to the side walls is perpendicular to $\bar{e}$. Consequently, there is no contribution to the integral from the side walls. It follows that

$$
\begin{equation*}
\iiint_{\mathcal{V}} \nabla \cdot\left(\frac{P^{2}}{\rho_{0} c} \bar{e}\right) d \mathcal{V}=\iint_{d \mathcal{A}(\bar{\xi})} \bar{e} \cdot\left(\frac{P^{2}}{\rho_{0} c} \bar{e}\right) d \mathcal{A}+\iint_{d \mathcal{A}\left(\bar{\xi}_{0}\right)}(-\bar{e}) \cdot\left(\frac{P^{2}}{\rho_{0} c} \bar{e}\right) d \mathcal{A}=0 \tag{11.3.21}
\end{equation*}
$$

The ray tube cross section is infinitesimal in extent, so the pressure may be considered to be constant along it. (The behavior at caustics will be addressed separately.) The direction $-\bar{e}$ in $d \mathcal{A}_{\mathrm{I}}$ corresponds to an inflow. Hence, the above relation states that the time-averaged power that flows into the ray tube at its beginning must equal the power that flows out at the end, that is

$$
\begin{equation*}
\frac{P(\bar{\xi})^{2}}{\rho_{0}(\bar{\xi}) c(\bar{\xi})} d \mathcal{A}(\bar{\xi})=\frac{P\left(\bar{\xi}_{0}\right)^{2}}{\rho_{0}\left(\bar{\xi}_{0}\right) c\left(\bar{\xi}_{0}\right)} d \mathcal{A}\left(\bar{\xi}_{0}\right) \tag{11.3.22}
\end{equation*}
$$

The explicit solution for the pressure amplitude at position $\bar{\xi}$ on a ray is

$$
\begin{equation*}
P(\bar{\xi})=\left[\frac{\rho_{0}(\bar{\xi}) c\left(\bar{\xi}^{\prime}\right)}{\rho_{0}\left(\bar{\xi}_{0}\right) c\left(\bar{\xi}_{0}\right)} \frac{d \mathcal{A}\left(\bar{\xi}_{0}\right)}{d \mathcal{A}(\bar{\xi})}\right]^{1 / 2} P\left(\bar{\xi}_{0}\right) \tag{11.3.23}
\end{equation*}
$$

Equations (11.3.23) and (11.2.23) are equivalent. Indeed, Eq. (11.3.22) could have been obtained from the principle that the time-averaged power that flows into the ray tube must equal the power that emerges. This was the basis for the analysis of vertically stratified media. In view of this observation, it is reasonable to ponder why the present derivation derived a transport equation, and then used it to determine the amplitude dependence. The answer is that the energy conservation argument does not hold if there is a flow in the ambient state, whereas a modified transport equation would be available.

The preceding developments indicate that we can determine the amplitude of the signal anywhere along a ray if we know the amplitude on that ray at some starting point. But to do so we must evaluate the ray tube area. With that determination as the objective, let us return to the situation addressed in Sect.11.2, specifically, an axisymmetric field in a depth-dependent fluid. The solution of the ray tracing equations gives position along a ray as a function of time. Different initial values of the grazing and azimuthal angles generate different rays, so position on a ray is a function of $t, \chi_{\mathrm{I}}$, and $\theta_{\mathrm{I}}$. However, by virtue of axisymmetry, the range $R$ and depth $z$ of a point depend only on $t$ and $\chi_{\mathrm{I}}$ because the azimuth angle is invariant along a ray. Rather than considering $t$ to be the independent variable, it is advantageous for this discussion to use the arclength $\ell$ measured along a ray from the starting position. This quantity is in a one-to-one correspondence to $t$. Indeed, the differential position increment in a time $t$ is $d \ell=c(\bar{\xi}) d t$, so we could determine $\ell$ by a direct integration along the ray path. With this change of variables, the transverse distance and depth are described functionally as $R\left(\ell, \chi_{\mathrm{I}}\right)$ and $z\left(\ell, \chi_{\mathrm{I}}\right)$. This is the situation depicted in Fig. 11.16. (The difference between this Figure and Fig. 11.11 is that $R$ previously was considered to depend explicitly on $z$.)

A ray tube cross section consists of all points at the same $t$. Because all rays within the tube are very close, the rays intersect the cross section at the same arclength $\ell$. Thus, the coordinates of the end point on all upper rays in Fig. 11.16 are $R\left(\ell, \chi_{\mathrm{I}}\right)$ and $z\left(\ell, \chi_{\mathrm{I}}\right)$, while the coordinates at the end of all lower rays are $R\left(\ell, \chi_{\mathrm{I}}+d \chi_{\mathrm{I}}\right)$ and $z\left(\ell, \chi_{\mathrm{I}}+d \chi_{\mathrm{I}}\right)$.


Fig. 11.16 Evaluation of ray tube area for the axisymmetric field radiated by a point source. The initial grazing angle is $\chi_{\mathrm{I}}$, the azimuthal angle is $\theta$, and $\ell$ is the arclength along a ray measured from the source

The Pythagorean theorem gives the distance in the vertical plane between these points,

$$
\begin{align*}
d h & =\left\{\left[R\left(\ell, \chi_{\mathrm{I}}+d \chi_{\mathrm{I}}\right)-R\left(\ell, \chi_{\mathrm{I}}\right)\right]^{2}+\left[z\left(\ell, \chi_{\mathrm{I}}+d \chi_{\mathrm{I}}\right)-z\left(\ell, \chi_{\mathrm{I}}\right)\right]^{2}\right\}^{1 / 2} \\
& =\left[\left(\frac{\partial R}{\partial \chi_{\mathrm{I}}}\right)^{2}+\left(\frac{\partial z}{\partial \chi_{\mathrm{I}}}\right)^{2}\right]^{1 / 2} d \chi_{\mathrm{I}} \tag{11.3.24}
\end{align*}
$$

The distance in the horizontal plane between the rays bounding the tube is $R d \theta_{\mathrm{I}}$. The corresponding cross-sectional area is

$$
\begin{equation*}
d \mathcal{A}(\bar{\xi})=R\left[\left(\frac{\partial R}{\partial \chi_{\mathrm{I}}}\right)^{2}+\left(\frac{\partial z}{\partial \chi_{\mathrm{I}}}\right)^{2}\right]^{1 / 2} d \chi_{\mathrm{I}} d \theta_{\mathrm{I}} \tag{11.3.25}
\end{equation*}
$$

This expression would be useful if we knew $R$ and $z$ as analytical functions, but numerical solution of the ray tracing equations gives a digitized representation. In that case if two adjacent rays $n$ and $n+1$ are launched at grazing angles that differ by a very small value $\Delta \chi_{\mathrm{I}}$, then the derivatives may be approximated as first-order finite differences,

$$
\begin{equation*}
\frac{\partial R}{\partial \chi_{\mathrm{I}}} \approx \frac{R_{n+1}-R_{n}}{\Delta \chi_{\mathrm{I}}}, \quad \frac{\partial z}{\partial \chi_{\mathrm{I}}} \approx \frac{z_{n+1}-z_{n}}{\Delta \chi_{\mathrm{I}}} \tag{11.3.26}
\end{equation*}
$$

where $R_{n}$ and $z_{n}$ are the coordinates on ray $n$ at the instant of interest, and $R_{n+1}$ and $z_{n+1}$ are the coordinates for ray $n+1$ at that instant.

A hint as to how to proceed when the field is not axisymmetric comes from the observation that the bracketed term in Eq.(11.3.25) describes $d h$. This distance is measured in the vertical plane between points on adjacent rays that are at the same arclength. Thus, $d h$ represents the magnitude of the change of position when
$\ell$ and $\theta_{\mathrm{I}}$ are held constant. An equivalent statement is that $d h$ is derived from the Jacobian of the transformation from the $(R, z, \theta)$ position variables to the $\left(\ell, \chi_{\mathrm{I}}, \theta_{\mathrm{I}}\right)$ variables of a ray path when $\theta_{\mathrm{I}}$ and $\ell$ are held constant. A description of ray tube area in terms of a Jacobian is the basic approach described by Jensen et al. ${ }^{9}$ for handling an arbitrary situation. Ultimately, evaluation of this quantity would be done via numerical methods because the ray paths are known as discrete data sets. Rather than following that approach, we will develop a formulation that relies solely on vector algebra.

In Fig. 11.17, a ray tube is formed by a selected ray, designated as number 1, and two adjacent rays, numbered 2 and 3 . The adjacent rays must be independent of ray 1. For example, if ray 1 corresponds to launch angles $\chi_{\mathrm{I}}$ and $\theta_{\mathrm{I}}$, then ray 2 could correspond to $\chi_{\mathrm{I}}+\Delta \chi_{\mathrm{I}}$ and $\theta_{\mathrm{I}}$, while ray 3 could correspond to $\chi_{\mathrm{I}}$ and $\theta_{\mathrm{I}}+\Delta \theta_{\mathrm{I}}$, where the angle increments are extremely small. This analysis of ray tube area requires that the evaluation of points on each ray use a common time base. (This would be the case if we prescribe the same output instants when the ray tracing equations are solved numerically.) This scheme allows us to identify points $\bar{\xi}_{1}, \bar{\xi}_{2}$, and $\bar{\xi}_{3}$ on the respective rays that correspond to any selected instant $t_{n}$. Therefore, these three points lie on the same wavefront $\tau=t_{n}$.

Fig. 11.17 Vectors used to evaluate the cross-sectional area of a ray tube


From $\bar{\xi}_{1}$ we construct relative position vectors to the other points,

$$
\begin{equation*}
\bar{h}_{2 / 1}=\bar{\xi}_{2}-\bar{\xi}_{1}, \quad \bar{h}_{3 / 1}=\bar{\xi}_{3}-\bar{\xi}_{1} \tag{11.3.27}
\end{equation*}
$$

The cross section area of this ray tube is $\Delta \mathcal{A}=\left|\bar{h}_{2 / 1}\right|\left|\bar{h}_{3 / 1}\right| \sin \left(\operatorname{angle}\left(\bar{h}_{2 / 1}, \bar{h}_{3 / 1}\right)\right)$. This is the magnitude of a scalar triple product, with the tangent vector as the third vector. Furthermore, although this area is finite, the fact that the rays correspond to initial conditions that are only slightly different allows us to consider the area to be an approximation of $d \mathcal{A}$, so we have found that

$$
\begin{equation*}
d \mathcal{A} \approx\left|\left(\bar{h}_{2 / 1} \times \bar{h}_{3 / 1}\right) \cdot \bar{e}_{1}\left(\bar{\xi}_{1}\right)\right| \equiv c\left(\bar{\xi}_{1}\right)\left|\left(\bar{h}_{2 / 1} \times \bar{h}_{3 / 1}\right) \cdot \bar{s}_{1}\left(\bar{\xi}_{1}\right)\right| \tag{11.3.28}
\end{equation*}
$$

[^53]If we wish to ascertain how the amplitude varies along a specific ray, which has been designated as ray 1 , we would employ the preceding description of $d \mathcal{A}$ to formulate Eq.(11.3.23) for each computed point on that ray. Alternatively, we can apply the computation of $d \mathcal{A}$ to the point on each ray at a common instant. The outcome of such an evaluation would be the distribution of pressure along a wavefront. Repetition of either calculation for all rays and all wavefronts would yield a map of the pressure field.

## Reflections and Caustics

The procedure we have developed entails computing a set of rays by solving the ray tracing equations. Each ray corresponds to a different initial condition at $t_{\mathrm{I}}=0$, which may be a different initial position $\bar{\xi}_{\mathrm{I}}$ and/or a different initial tangent direction $\bar{e}_{\mathrm{I}}$. The differential equation solution marches forward incrementally. Upon reaching the farthest location, the solver is halted. After all rays have been computed, the properties of the rays may be used to determine how the pressure varies along any ray, provided that the computation has described a sufficient number of rays to approximate ray tubes.

The evaluation of each ray may be carried without interruption, unless a ray encounters a boundary. In that case, the ray is reflected. In view of the locally planar nature of a signal in the geometrical acoustics approximation, the rules for reflection of a plane wave that is incident on an infinite plane may be employed here. In Fig. 11.18, a ray is incident on a solid boundary at $\bar{\xi}_{B}$. The normal to the surface, pointing into the fluid, at this location is $\bar{n}\left(\bar{\xi}_{B}\right)$, and $\bar{e}\left(\bar{\xi}_{B^{-}}\right)$is the tangent to the ray path immediately before it touches the boundary. The tangent to the reflected ray immediately after it leaves the surface is $\bar{e}\left(\bar{\xi}_{B^{+}}\right)$. The angle of incidence $\psi_{B}$ equals the angle of reflection.

Fig. 11.18 Unit vectors describing the incidence and reflection of a ray at position $\bar{\xi}_{B}$ in a solid surface


We wish to describe $\bar{e}\left(\bar{\xi}_{B^{+}}\right)$in a manner that does not require defining a coordinate system. To that end, let $\bar{e}_{t}\left(\bar{\xi}_{B}\right)$ denotes the tangent to the surface at the point of incidence in the plane formed by $\bar{e}\left(\bar{\xi}_{B^{-}}\right)$and $\bar{n}\left(\bar{\xi}_{B}\right)$. It follows that the component representation of the tangent directions is

$$
\begin{equation*}
\bar{e}\left(\bar{\xi}_{B^{-}}\right)=\bar{e}_{t}\left(\bar{\xi}_{B}\right) \sin \psi_{B}-\bar{n}\left(\bar{\xi}_{B}\right) \cos \psi_{B}, \quad \bar{e}\left(\bar{\xi}_{B^{+}}\right)=\bar{e}_{t}\left(\bar{\xi}_{B}\right) \sin \psi_{B}+\bar{n}\left(\bar{\xi}_{B}\right) \cos \psi_{B} \tag{11.3.29}
\end{equation*}
$$

Subtracting the first equation from the second yields the desired relation. The angle of incidence may be found from a dot product, so we have

$$
\begin{equation*}
\bar{e}\left(\bar{\xi}_{B^{+}}\right)=\bar{e}\left(\bar{\xi}_{B^{-}}\right)+2 \bar{n}\left(\bar{\xi}_{B}\right) \cos \psi_{B}, \quad \psi_{B}=\cos ^{-1}\left(-\bar{e}\left(\bar{\xi}_{B^{-}}\right) \cdot \bar{n}\left(\bar{\xi}_{B}\right)\right) \tag{11.3.30}
\end{equation*}
$$

Subsequent to reflection, solution of the ray tracing equations may resume. The initial conditions for that computation would be the values of $\bar{\xi}_{B}$ and $\bar{e}\left(\bar{\xi}_{B^{+}}\right)$at $t_{B}$.

The computation of the amplitude variation along a ray also must account for a reflection. Because the angles of incidence and reflection are equal, the ray tube area immediately prior to incidence is the same as the area after reflection. Hence, the only reason the amplitude might be altered by the reflection is that the reflection coefficient is not unity. Suppose the surface is locally reacting with impedance $Z$. Then, the reflection coefficient is

$$
\begin{equation*}
\mathcal{R}=\frac{Z\left(\bar{\xi}_{B}\right)-\rho_{0} c / \cos \psi_{B}}{Z\left(\bar{\xi}_{B}\right)+\rho_{0} c / \cos \psi_{B}} \tag{11.3.31}
\end{equation*}
$$

Neither the local impedance nor the angle of incidence is required to be constant along the surface. If that is the situation, some rays will be attenuated more than others as a result of reflection. However, this effect often is ignored by taking the surface to be rigid, $\mathcal{R}=1$, based on the fact that at very high frequencies most materials have a large impedance. In any case, the amplitude immediately after reflection is $\mathcal{R}$ times the amplitude before. Using the properties at the reflection point as the input for the reflected ray tube leads to a description of the reflected signal as

$$
\begin{equation*}
p(\bar{\xi}, t)=\operatorname{Re}\left(\left[\frac{\rho_{0}(\bar{\xi}) c(\bar{\xi})}{\rho_{0}\left(\bar{\xi}_{B}\right) c\left(\bar{\xi}_{B}\right)} \frac{d \mathcal{A}\left(\bar{\xi}_{B}\right)}{d \mathcal{A}(\bar{\xi})}\right]^{1 / 2} \mathcal{R} P\left(\bar{\xi}_{B^{-}}\right) e^{i \omega(t-\tau(\bar{x}))}\right) \tag{11.3.32}
\end{equation*}
$$

The occurrence of a caustic is the consequence of the global behavior of a family of rays, rather than a special feature of a single ray. Thus, there is no need to interrupt solution of the ray tracing equations upon arrival at a caustic. (Indeed, one would be unlikely to recognize the existence of a caustic until the full set of rays has been computed.) The occurrence of a caustic does alter the amplitude dependence. We shall merely describe how to account for that effect because the analysis is beyond the present scope. The interested reader is referred to Pierce, ${ }^{10}$ which also contains references to the original analyses.

Mathematical analysis would reveal that a harmonic signal undergoes a $\pi / 2$ phase delay as it propagates along its ray through a caustic. Specifically, if $\ell_{A}$ and $\ell_{B}$, respectively, are the arclengths to points on a ray prior to, and following, the caustic at $\ell_{\text {caustic }}$, then we have

$$
\begin{gather*}
p\left(\bar{\xi}_{B}, t\right)=\operatorname{Re}\left(\left[\frac{c\left(\bar{\xi}_{B}\right)}{c\left(\bar{\xi}_{A}\right)} \frac{d \mathcal{A}\left(\bar{\xi}_{A}\right)}{d \mathcal{A}\left(\bar{\xi}_{B}\right)}\right]^{1 / 2} P\left(\bar{\xi}_{A}\right) e^{i \omega(t-\tau(\bar{x}))}\right) \text { if } \ell_{B}<\ell_{\text {caustic }} \\
p\left(\bar{\xi}_{B}, t\right)=\operatorname{Re}\left(\left[\frac{c\left(\bar{\xi}_{B}\right)}{c\left(\bar{\xi}_{A}\right)} \frac{d \mathcal{A}\left(\bar{\xi}_{A}\right)}{d \mathcal{A}\left(\bar{\xi}_{B}\right)}\right]^{1 / 2} P\left(\bar{\xi}_{A}\right) e^{-i \pi / 2} e^{i \omega(t-\tau(\bar{x}))}\right) \text { if } \ell_{B}>\ell_{\text {caustic }} \tag{11.3.33}
\end{gather*}
$$

[^54]If it should be that two caustics are situated on a ray between $\bar{\xi}_{A}$ and $\bar{\xi}_{B}$, then the phase delay would be $2(\pi / 2)$, and so on. Note that the ray tube area decreases to zero at a caustic, which means that the linear acoustics approximation requires correction. In the vicinity of the caustic, the amplitude will be greatly enhanced.

A heuristic explanation of this phase shift may be obtained by considering the ray tube area beyond a caustic to be negative. This view stems from the description of ray tube area as a proportionality to the Jacobian. The rays cross over as they pass the caustic. The Jacobian is positive ahead of the caustic, zero at the caustic, and negative beyond it. The amplitude is proportional to the square root of the area, and the square root of a negative number gives a $\pm i \equiv \exp ( \pm i \pi / 2)$. The root that applies is $-i$ for a delay of the signal. This aspect may be viewed as a manifestation of the principle of causality. A factor $+i$ as the signal passes through a caustic corresponds to a phase lead. If that were the case, it would mean that a feature is observed at a location infinitesimally beyond the caustic prior to its observation at a location that infinitesimally precedes the caustic.

EXAMPLE 11.5 Air is contained within a rectangular waveguide whose side walls are composed of tempered glass, which is well approximated as being rigid. As sketched below, circular plates of radius $a$ are attached to the horizontal walls. The plates are cooled, thereby altering the sound speed in the region between them. Because the depth $H$ is small and the walls are rigid, it is reasonable to consider the sound speed to vary only in the horizontal directions. Hence, the cooled region may be taken to constitute a vertical cylinder of lower sound speed. The specific dependence is $c=c_{0}+\left(c_{1}-c_{0}\right)\left(1-R^{2} / a^{2}\right)$, where $R$ is the transverse distance from the centerline connecting the plates. Outside this region, the sound speed is $c_{0}$. The wall at $x=0$ generates a 200 kHz harmonic plane wave, and the wall at $x=L$ is actively controlled to eliminate reflections. Draw the rays and wavefronts. Then determine the dependence of pressure amplitude on distance along the wavefront that touches the front of the cylindrical region, where $x=2.5 a$. Parameters for the evaluation are $c_{0}=330 \mathrm{~m} / \mathrm{s}, c_{1}=280 \mathrm{~m} / \mathrm{s}$, and $a=250 \mathrm{~mm}$.


## Significance

Application of numerical methods to solve the ray tracing and transport equations is a primary thrust, and the results will greatly enhance our understanding of the behavior of rays and wavefronts, including a much deeper insight to the behavior near caustics.

## Solution

The sound speed depends on the transverse distance $R$, but the waveguide is best described in terms of Cartesian coordinates. Because the sound speed is independent of the vertical position $y$ and the top and bottom walls are rigid, the pressure field will be independent of $y$, so $\bar{s}$ and $\bar{\xi}$ may be described solely in terms of $x$ and $z$ components. The column vector of variables for a numerical solution is

$$
\{X\}=\left[\begin{array}{llll}
s_{x} & s_{z} & \xi_{x} & \xi_{z}
\end{array}\right]^{\mathrm{T}}
$$

The corresponding vector driving $d\{X\} / d t$ is

$$
\{F\}=\left[\begin{array}{ll}
\left.-\frac{1}{c}\left[\begin{array}{ll}
\frac{\partial c}{\partial \xi_{x}} & \frac{\partial c}{\partial \xi_{z}}
\end{array}\right] \quad c^{2}\left[\begin{array}{ll}
s_{x} & s_{z}
\end{array}\right]\right]^{\mathrm{T}} .4{ }^{\mathrm{T}} .
\end{array}\right.
$$

The transverse distance from the vertical centerline of the cooling plates to $\bar{\xi}$ is $R=\left|\bar{\xi}-1.5 a \bar{e}_{x}\right|$, which is used to evaluate the given expression for $c(\bar{\xi})$. The corresponding gradient of $c$ is $\nabla c=(d c / d R) \nabla R$. We know that $\nabla R=\bar{e}_{R}$, where $\bar{e}_{R}$ is the unit vector from the center to point $\bar{\xi}$. Thus,

$$
\nabla c=\frac{d c}{d R} \bar{e}_{R}=\left\{\begin{array}{l}
-\frac{2 R}{a^{2}}\left(c_{1}-c_{0}\right) \bar{e}_{R} \text { if } R<a \\
0 \text { if } R \geq a
\end{array}\right.
$$

This gradient must be described in term of $x$ and $z$ components, with $\xi_{x}$ and $\xi_{z}$ as the variables. The required expressions are obtained by a transformation to Cartesian coordinates, which sets

$$
\bar{\xi}=\xi_{x} \bar{e}_{x}+\xi_{z} \bar{e}_{z}, \quad R=\left|\bar{\xi}-1.5 a \bar{e}_{x}\right|, \quad \bar{e}_{r}=\frac{\bar{\xi}-1.5 a \bar{e}_{x}}{R}
$$

The numerical solver routine must be selected. Matlab® offers several. The one that was used to generate the data is ODE45, which implements a fourth/fifth-order Runge-Kutta algorithm. The input arguments for this routine are the name of a function $F$ _rhs that will evaluate $\{F\}$, followed by a column or row vector $t$ _span holding $t$ values at which the solution will be output. The next input argument is the initial value of $\{X\}$, which will be addressed below. The last input argument for ODE45 is an optional set of parameters setting error limits. The results here were obtained by setting the relative error tolerance to $10^{-6}$.

There are two alternative procedures for stepping through time with ODE45. If t_span holds all of the time instants, the output from a single call to ODE45 would be a rectangular array, which we will denote as [ $X_{\text {out }}$ ]. The $n$th row of $\left[X_{\text {out }}\right]$ would be the solution $\{X\}$ at time $t_{n}$. The alternative procedure explicitly steps through time by setting $t_{-}$span as a two-element vector $\left[\begin{array}{ll}t_{n-1} & t_{n}\end{array}\right]$. The value of $\{X\}$ at $t_{n}$ is the last row of $\left[X_{\text {out }}\right]$ in that case. (The other rows are $\{X\}$ values at intermediate instants listed in $\left\{t_{\text {out }}\right\}$.) The time history for all variables in this procedure may be constructed by adjoining the transpose of the column vectors $\{X\}$. In either treatment, the time values at which data is output are stored in a column vector $\{t\}$. Clearly, the singlecall procedure is easier to implement, and it was found to require much less time to execute. However, the explicit stepping procedure was found to be significantly more accurate.

The numerical integration must be performed for each ray. This was done within a loop over the ray number. The solution for all rays should use a common base of $t$ values. Doing so aligns the data such that the computed $\bar{\xi}$ values for all rays lie on a common set of wavefronts. The time increment for the evaluation was taken to be the time required to travel a distance $a / 25$ at speed $c_{0}$, and the maximum time was set as the time to travel $6 a$ at speed $c_{0}$. Thus, the number of $t_{n}$ values is $N=\left(6 a / c_{0}\right) /\left(a / 25 c_{0}\right)=150$. Each pass through the ray number loop corresponds to a ray that departs from the plane $x=0$ at an incremented distance $z_{j}$ in the $z$ direction. Each ray initially is a straight line, so $\bar{e}=\bar{e}_{x}$ at the start, which is defined to be $t=0$. Let $J$ be the number of rays. The diameter of the cooled region is $2 a$, so the rays were initialized at a uniform spacing to cover a width of $3 a$, which gives a separation of $\Delta=3 a /(J-1)$ between rays. The corresponding initial conditions for the solution for ray $j$ were $s_{x}=1 / c_{0}, s_{z}=0, \xi_{x}=0, \xi_{z}=-1.5 a+\Delta(j-1)$.

A single pass through the outer loop over ray numbers gives $\left[X_{\text {out }}\right]$ for ray $j$. The four rows of $\left[X_{\text {out }}\right]$ consist, respectively, of $s_{x}, s_{z}, \xi_{x}, \xi_{z}$ at each $t_{j}$. The row of $\xi_{x}$ data was stored as column $j$ of a rectangular array $\left[\xi_{x}\right]$, and the $\xi_{z}$ row was likewise stored in column $j$ of $\left[\xi_{x}\right]$. The eikonal error metric for each ray, numbered $j$, was obtained by evaluating $c^{2}\left[\left(s_{x}\right)^{2}+\left(s_{z}\right)^{2}\right]-1$ for each instant $t_{n}$, and then storing the error values at all instants as column $j$ of the rectangular array $\left[\varepsilon_{\text {eik }}\right]$. The largest eikonal errors were found to occur along rays 21 and 25 . Figure 2 displays the variation of this quantity along the respective rays. The largest error for any ray was found to be less than $0.08 \%$.


Figure 2.

When the data is stored as described above, a ray may be drawn by plotting a column of $\left[\xi_{x}\right]$ against the corresponding column of $\left[\xi_{y}\right]$, while wavefronts are obtained by plotting a row of $\left[\xi_{x}\right]$ against the corresponding row of $\left[\xi_{y}\right]$. The rays and wavefronts appear in Fig. 3. For the sake of clarity, only every tenth of the evaluated wavefronts is shown there. There are many interesting facets to this picture. The rays that pass through the cooled region are bent inward toward the $x$-axis. This is consistent with the property observed for vertical stratification. It results from a wavefront being retarded where the sound speed is lower, thereby having the effect of turning the wavefront. There appears to be regions where there is no ray, but these are not shadow zones. Rather the rays are sparse there. If we were to greatly increase the number of computed rays, we would see that rays do pass through those regions. Regions where the rays are sparse correspond to low, but not zero, pressure.


Figure 3.

Figure 3 shows that there is a region in which many rays converge. Figure 4 zooms in on this region. The concentration is not a focus because the rays do not intersect at a single point. Another feature leading to the same conclusion is that the wavefronts beyond the convergence are not circles, as they would be if the rays were radiated from a focus. Rather, the concentration of rays is a caustic. Inspection of a region beyond the arrete shows that at any point there are three rays. Two arrive at comparatively large angles from either side of the $x$-axis, whereas the third comes at a small angle from the opposite side of the $x$-axis. This results in folding of a wavefront. The part closer to the arrete is a curved arc. It is the locus of constant $\tau$ values for the shallow rays. The portions of a folded wavefront that intersect at the $x$-axis correspond to the large angle rays.


Figure 4.

The location of the caustic can be shifted by adjusting the value of $c_{1}$ with $c_{0}$ fixed. Decreasing $c_{1}$ increases the heterogeneity, thereby increasing the curvature of the rays within the cooled region. This brings the caustic closer to the origin. Indeed, if $c_{1}$ is sufficiently large, the caustic will occur inside that region.

It still remains to evaluate the pressure along a wavefront. For this analysis, there are several items we must determine. First, we must determine where the computed ray paths intersect the wavefront that was stipulated. The point where this wavefront tangentially touches the cooled region is $\xi_{x}=2.5 a=0.625 \mathrm{~m}, \xi_{z}=0$. The ray passing through this point is the straight one along the center line. Its ray number is $j=(J+1) / 2=31$. None of the computed $\xi_{x}$ values for this ray exactly equals 0.625 m . We can use interpolation to reach the specified wavefront. First, we scan the $\xi_{x}$ values for ray 31 to identify the points that bracket $\xi_{x}=0.625 \mathrm{~m}$. These are $\left(\xi_{x}\right)_{69,31}=0.6225$ and $\left(\xi_{x}\right)_{70,31}=0.6325$. A linear interpolation between these points should give $\left(\xi_{x}\right)_{\mathrm{wff}, 31}=0.625 \mathrm{~m}$, that is,

$$
0.625=0.6225+\beta(0.6325-0.6225) \Longrightarrow \beta=0.25
$$

The interpolation fraction $\beta$ is used to find the value of $\tau$ for the wavefront, according to

$$
\tau_{\mathrm{wf}}=\tau_{69}+\beta\left(\tau_{70}-\tau_{69}\right)=2.0681 \mathrm{~ms}
$$

This interpolation also is employed to determine the $x$ and $y$ coordinates at which any ray $j$ intersects the wavefront, according to

$$
\left\{\begin{array}{l}
\left(\xi_{x}\right)_{\mathrm{wf}, j} \\
\left(\xi_{y}\right)_{\mathrm{wf}, j}
\end{array}\right\}=\left\{\begin{array}{l}
\left(\xi_{x}\right)_{69, j} \\
\left(\xi_{y}\right)_{70, j}
\end{array}\right\}+\beta\left\{\begin{array}{l}
\left(\xi_{x}\right)_{70, j}-\left(\xi_{x}\right)_{69, j} \\
\left(\xi_{y}\right)_{70, j}-\left(\xi_{y}\right)_{70, j}
\end{array}\right\}, \quad j=1, \ldots, J
$$

Now that the ray-wavefront intersections have been located, the next task is to determine the ray tube areas. Reference to Fig. 5 will assist following this development. The depth perpendicular to ray diagrams is constant, and we are only interested in the ratio of areas. Therefore, an area ratio reduces to the ratio of distances between intersection points. Let $h_{\mathrm{wf}, j}$ be the distance between the points at which rays $j$ and $j+1$ intersect the wavefront. These distances are

$$
h_{\mathrm{wf}, j}=\left|\bar{\xi}_{\mathrm{wf}, j+1}-\bar{\xi}_{\mathrm{wf}, j}\right|, \quad j=1, \ldots, J-1
$$

The rays are equally spaced at distance $\Delta$ when they depart from $x=0$, so $h_{1, j}=$ $\Delta$ for all $j$.


Figure 5.

The last quantity required to evaluate the pressure amplitude according to Eq. (11.3.23) is the characteristic impedance. The stipulated wavefront does not lie in the cooled region, so the ratio of $\rho_{0} c$ along it to the $\rho_{0} c$ value at the initial point is unity. Nevertheless, we shall incorporate this effect into the computation because doing so would allow us to probe inside the cooled region. The value of $\rho_{0}$ is not specified for any location, but we can deduce it from the given ambient properties. In the ambient state, the fluid everywhere is quiescent, so there can be no pressure gradient. It follows that the pressure within the cooled region must be the same as it is outside this region. We shall assume that the ambient pressure is atmospheric, so $P_{0}=1 \mathrm{~atm}=101.32 \mathrm{kPa}$. We know that $c=\left(\gamma P_{0} / \rho_{0}\right)^{1 / 2}$ for an ideal fluid, with $\gamma=1.4$ for air, and the value of $c(\bar{\xi})$ at any location is given. Solving this relation for the density gives $\rho_{0}(\bar{\xi})=\gamma P_{0} / c(\bar{\xi})^{2}$.

The ray tube numbered $j$ covers the distance between $\bar{\xi}_{\mathrm{wf}, j}$ and $\bar{\xi}_{\mathrm{wf}, j+1}$. It is reasonable to use the impedance midway between those points to formulate Eq. (11.3.23). These values are $\left(\rho_{0} c\right)_{\mathrm{av}, j} \equiv \rho_{0}\left(\bar{\xi}_{\mathrm{av}, j}\right) c\left(\bar{\xi}_{\mathrm{av}, j}\right)$, where $\bar{\xi}_{\mathrm{av}, j}$ is the midpoint position,

$$
\bar{\xi}_{\mathrm{av}, j}=\frac{1}{2}\left(\bar{\xi}_{\mathrm{wf}, j}+\bar{\xi}_{\mathrm{wf}, j+1}\right)
$$

The pressure at the midpoint corresponding to these parameters is indicated by Eq.(11.3.23) to be

$$
\frac{P_{\mathrm{wf}, j}}{P_{\mathrm{I}}}=\left(\frac{\left(\rho_{0} c\right)_{\mathrm{av}, j}}{\left(\rho_{0} c\right)_{x=0}}\right)^{1 / 2}\left(\frac{\Delta}{h_{\mathrm{wf}, j}}\right)^{1 / 2}
$$

where $P_{\mathrm{I}}$ is the pressure amplitude in the plane wave at $x=0$.
These pressure ratios are associated with the midpoint between rays. To plot them, we will use as the abscissa the distance along the wavefront measured from the point where the wavefront intersects the first ray. This distance, which we shall denote as $\eta_{\mathrm{wf}, j}$, is shown in Fig. 5 to be approximately the accumulation of the ray tube widths preceding it, according to.

$$
\eta_{\mathrm{wf}, 1}=\frac{1}{2} h_{\mathrm{wf}, 1}, \quad \eta_{j}=\frac{1}{2} h_{\mathrm{wf}, j}+\sum_{n=1}^{j-1} h_{\mathrm{wf}, n}
$$

Figure 6 plots these the pressure ratio against distance along the wavefront, with the abscissa shifted such that $\eta=0$ at the centerline. The number of rays for this graph was raised to 121 to give the graphs a less jagged appearance. The wavefront that was specified contains the point $x=0.625 \mathrm{~m}, y=0$. Its equation is $\tau=2.0681 \mathrm{~ms}$. The pressure distribution along two other wavefronts also appear in this graph. The point $x=0.5 \mathrm{~m}, y=0$ lies on the wavefront $\tau=1.6633 \mathrm{~ms}$. This wavefront passes through the cooled region. The third wavefront is $\tau=2.4469 \mathrm{~ms}$. It contains the point $x=0.75 \mathrm{~m}, y=0$, which places it slightly to the left of the arrete. Reference to Fig. 3 shows that the intervals where $|P| / P_{\mathrm{I}}$ is less than one correspond to regions where the rays are sparse, while the rays are closely spaced in regions where the pressure ratio is enhanced. This enhancement becomes more prominent as the wavefront comes closer to the caustic.


Figure 6.

## Evaluation of the Signal at a Field Point - Eigenrays

The operations described thus far entail evaluating ray paths by solving the ray tracing equations. The pressure along a ray is then determined by solving the transport equation. In this procedure, we cannot specify a priori the points where the ray falls. This leads to the question of how can we find the pressure at a specific point? Stated differently, the question is how can we determine the ray path that connects a source and a specific field point?

The situation is depicted in Fig. 11.19 where the source is at $\bar{\xi}_{\mathrm{I}}$, and the receiver's location is the designated field point $\bar{\xi}_{f}$. Only direct and surface-reflected rays are considered. It is permissible to ignore bottom-reflected rays if the water is very deep compared to the depths of the source and field point, because signals that propagate along bottom-reflected rays to $\bar{\xi}_{f}$ will have traveled over very large arclengths. Unless a caustic occurs, the ray tube area of this set of rays will be greatly enlarged, so the amplitude will be greatly reduced. (Bottom reflections are an important aspect of propagation in a shallow channel.)


Fig. 11.19 Identification of eigenray connecting a point source and a selected field point
Contributors to the signal at $\bar{\xi}_{f}$ are a ray that arrives directly from the source, and a ray that arrives after it is reflected from the free surface. Both are referred to as eigenrays, which is the term given generally to a ray that connects the source and a selected field point. Solution of the ray tracing equations presumably has given the two families of rays in Fig. 11.19, where each ray corresponds to a different grazing angle at the source. How can we identify the eigenrays given that none of the rays in the Figure actually intersects $\bar{\xi}_{f}$ ? This is not a trivial questions, and it is much more difficult if some rays undergo many surface and bottom reflections.

The simplest approach is to use interpolation. To do so, we first identify the pair of direct rays $n$ and $n+1$ that bracket $\bar{\xi}_{f}$, that is, the rays that come closest to $\bar{\xi}_{f}$ above
and below. Let $R_{f}$ be the horizontal range to $\bar{\xi}_{f}$, and let $z_{n}$ and $z_{n+1}$ be the depths of the points on the respective rays that are at $R_{f}$. A linear interpolation of the initial grazing angles gives

$$
\begin{equation*}
\frac{\left(\chi_{\mathrm{I}}\right)_{f}-\left(\chi_{\mathrm{I}}\right)_{n}}{z_{f}-z_{n}}=\frac{\left(\chi_{\mathrm{I}}\right)_{n+1}-\left(\chi_{\mathrm{I}}\right)_{n}}{z_{n+1}-z_{n}} \tag{11.3.35}
\end{equation*}
$$

The value of $\left(\chi_{\mathrm{I}}\right)_{f}$ is used to find the eigenray. To do so, the ray tracing equations are solved with $\left(\chi_{\mathrm{I}}\right)_{f}$ as the initial grazing angle. The eigenray for surface reflection is found by repeating these steps using the rays in the reflected set that bracket $\bar{\xi}_{f}$.

The total pressure at $\bar{\xi}_{f}$ is the sum of the contributions from all eigenrays connecting the source and field points. In the case of Fig. 11.19, let $\tau_{d}$ and $\tau_{r}$ be the values of the wavefront functions for the direct and reflected eigenrays, respectively. Also, let $P_{d}$ and $P_{r}$ be the corresponding amplitudes. Both amplitudes are obtained by evaluating the cross-sectional areas of the respective ray tubes, with the appropriate corrections for reflection coefficients and phase delays at caustics. Then the total pressure at $\bar{\xi}_{f}$ is

$$
\begin{equation*}
p\left(\bar{\xi}_{f}, t\right)=\operatorname{Re}\left[P_{d} e^{i \omega\left(t-\tau_{d}\right)}+P_{r} e^{i \omega\left(t-\tau_{r}\right)}\right] \tag{11.3.36}
\end{equation*}
$$

If more than two eigenrays intersect the selected field point, than the contribution of each is additive.

EXAMPLE 11.6 A radially symmetric point source is situated at depth $H$ below the free surface of a deep body of water. Density and sound speed variations are negligible. Use geometrical acoustics to derive an expression for the pressure at range $R_{F}$ and depth $z_{F}$. Sound speed variations over this range are negligible. Compare that result to the result of an analysis that uses the method of images.

## Significance

The main feature of the analysis is that it explains how the Jacobian is used to evaluate analytically the ray tube area. The juxtaposition of geometrical acoustics and the method of images illustrates their different perspectives.

## Solution

The fluid is homogeneous, so all rays are straight lines. There are only two eigenrays because reflections from the bottom are deemed to be negligible. The direct eigenray is the line from the source to the field point, and the reflected eigenray intersects the surface somewhere between the source and the receiver. The arrangement is described in Fig. 1.


Figure 1.

A geometrical acoustics analysis begins with determination of the ray paths. Because they are formed from straight lines, we may use geometry for the determination, rather than solving the ray tracing equations. The angle of incidence equals the angle of reflection. Consequently, right triangles $I A B$ and $B C F$ are similar. Equating the ratio of sides yields an expression for the horizontal distance $s$ to the reflection point,

$$
\frac{s}{H}=\frac{R_{F}-s}{z_{F}} \Longrightarrow s=\frac{H R_{F}}{z_{F}+H}
$$

We must evaluate ray tube areas for each eigenray. We could do so by considering each eigenray to be one limit of its ray tube with rays at initial grazing angles $\chi_{d}+d \chi$ and $\chi_{r}+d \chi$ as the other limit. However, doing so would be tedious. An analytical approach employs the Jacobian in Eq.(11.3.25). This formulation requires description of the $(R, z)$ coordinates of a field in terms of the initial grazing angle for the ray path through that point and the distance $\ell$ along that ray path to the field point. The source depth $H$ is fixed in that evaluation

For the direct path, the arclength is $\ell_{d}$, and the initial grazing angle is $\chi_{d}$. In terms of these variables, the coordinates of the field point $F$ are

$$
R_{F}=\ell_{d} \cos \chi_{d}, \quad z_{F}=H+\ell_{d} \sin \chi_{d}
$$

We differentiate these expressions with respect to $\chi_{d}$, and substitute the results into Eq. (11.3.25). This gives

$$
\begin{equation*}
d \mathcal{A}_{d}=R_{F}\left[\left(\frac{\partial R_{F}}{\partial \chi_{d}}\right)^{2}+\left(\frac{\partial z_{F}}{\partial \chi_{d}}\right)^{2}\right]^{1 / 2} d \chi_{d} d \theta_{d}=\left(\ell_{d}\right)^{2}\left(\cos \chi_{d}\right) d \chi_{d} d \theta \tag{1}
\end{equation*}
$$

The presence of a reflection complicates the evaluation of ray tube area for reflected rays. We seek a description of $R_{F}$ and $z_{F}$ in terms of $\chi_{r}$ and the total arclength along the reflected ray. Figure 1 indicates that this arclength is $\ell_{r}=\ell_{1}+\ell_{2}$. Because $H$ is fixed, the distance $\ell_{1}$ depends solely on $\chi_{r}$. Using this relation to form $\ell_{2}=\ell_{r}-\ell_{1}$ gives $\ell_{2}$ as a function of $\chi_{r}$ and $\ell_{r}$. We employ this relation to describe $R_{F}$ and $z_{F}$ for the reflected ray,

$$
\begin{gathered}
\ell_{1}=\frac{H}{\sin \chi_{r}}, \quad \ell_{2}=\ell_{r}-\ell_{1}=\ell_{r}-\frac{H}{\sin \chi_{r}} \\
R_{F}=\ell_{1} \cos \chi_{r}+\ell_{2} \cos \chi_{r}=\ell_{r} \cos \chi_{r} \\
z_{F}=\ell_{2} \sin \chi_{r}=\ell_{r} \sin \chi_{r}-H
\end{gathered}
$$

The ray tube area obtained from Eq. (11.3.25) is

$$
\begin{align*}
d \mathcal{A}_{r} & =R_{F}\left[\left(\frac{\partial R_{F}}{\partial \chi_{r}}\right)^{2}+\left(\frac{\partial z_{F}}{\partial \chi_{r}}\right)^{2}\right]^{1 / 2} d \chi_{\mathrm{I}} d \theta_{\mathrm{I}}  \tag{2}\\
& =R_{F}\left[\left(\ell_{r} \sin \chi_{r}\right)^{2}+\left(\ell_{r} \cos \chi_{r}\right)^{2}\right]^{1 / 2} d \chi_{r} d \theta=\left(\ell_{r}\right)^{2}\left(\cos \chi_{r}\right) d \chi_{r} d \theta
\end{align*}
$$

Equation(11.3.22) equates twice the intensity at an arbitrary location to the value at the start of the ray tube. In the present situation, the input to both ray tubes is a radially symmetric point source. The time-averaged intensity is the source's radiated power divided by the surface area of a surrounding sphere. For both ray tubes, this sphere has a very small radius $\varepsilon$. Therefore, for both the direct and reflected eigenvalues the starting value is

$$
\begin{equation*}
\left(\frac{P^{2}}{\rho_{0} c}\right)_{\mathrm{I}}=\frac{2 \mathcal{P}}{4 \pi \varepsilon^{2}} \tag{3}
\end{equation*}
$$

The cross-sectional area for both ray tubes at the input end consists of a small patch of the surrounding sphere at polar angle $\pi / 2-\chi$ covering a polar angle $d \chi$ and azimuthal angle $d \theta$. Therefore, we have

$$
\begin{equation*}
\left(d \mathcal{A}_{\mathrm{I}}\right)_{d}=\left(\varepsilon d \chi_{d}\right)\left(\varepsilon \cos \chi_{d}\right) d \theta, \quad\left(d \mathcal{A}_{\mathrm{I}}\right)_{r}=\left(\varepsilon d \chi_{r}\right)\left(\varepsilon \cos \chi_{r}\right) d \theta \tag{4}
\end{equation*}
$$

Application of Eq. (11.3.22) in the case where the input to a ray tube is a radially symmetric point source leads to

$$
\frac{P_{d}^{2}}{\rho_{0} c} d \mathcal{A}_{d}=\left(\frac{P^{2}}{\rho_{0} c}\right)_{\mathrm{I}}\left(d \mathcal{A}_{\mathrm{I}}\right)_{d} \text { and } \frac{P_{r}^{2}}{\rho_{0} c} d \mathcal{A}_{r}=\left(\frac{P^{2}}{\rho_{0} c}\right)_{\mathrm{I}}\left(d \mathcal{A}_{\mathrm{I}}\right)_{r}
$$

Substitution of Eqs. (1)-(4) into these relations gives

$$
\begin{gathered}
\frac{P_{d}^{2}}{\rho_{0} c} \ell_{d}^{2}\left(\cos \chi_{d}\right) d \chi_{\mathrm{d}} d \theta=\left(\frac{2 \mathcal{P}}{4 \pi \varepsilon^{2}}\right) \varepsilon^{2}\left(\cos \chi_{d}\right) d \chi_{\mathrm{d}} d \theta \\
\frac{P_{r}^{2}}{\rho_{0} c}\left(\ell_{r}\right)^{2}\left(\cos \chi_{r}\right) d \chi_{r} d \theta=\left(\frac{2 \mathcal{P}}{4 \pi \varepsilon^{2}}\right) \varepsilon^{2}\left(\cos \chi_{r}\right) d \chi_{r} d \theta \\
P_{d}=\frac{1}{\ell_{d}}\left(\frac{\rho_{0} c \mathcal{P}}{2 \pi}\right)^{1 / 2}
\end{gathered}
$$

From this we find that the pressure amplitudes are

$$
\begin{equation*}
P_{d}=\frac{1}{\ell_{d}}\left(\frac{\rho_{0} c \mathcal{P}}{2 \pi}\right)^{1 / 2}, \quad P_{r}=\frac{1}{\ell_{r}}\left(\frac{\rho_{0} c \mathcal{P}}{2 \pi}\right)^{1 / 2} \tag{5}
\end{equation*}
$$

For a wave that departs from the source at zero phase and propagates at constant phase speed $c$ along a ray, the phase variable is $\tau(\bar{x})=\ell / c$. The reflection coefficient at the free surface may be taken to be -1 . Adding the signal received at the field point from each eigenray therefore gives

$$
\begin{aligned}
p\left(\bar{\xi}_{f}, t\right) & =\operatorname{Re}\left[P_{d} e^{i \omega\left(t-\ell_{d} / c\right)}+P_{r}(-1) e^{i \omega\left(t-\ell_{r} / c\right)}\right]^{1 / 2} \\
& =\left(\frac{c \mathcal{P}}{2 \pi}\right)^{1 / 2} \operatorname{Re}\left[\left(\frac{e^{-i \omega \ell_{d} / c}}{\ell_{d}}-\frac{e^{-i \omega \ell_{r} / c}}{\ell_{r}}\right) e^{i \omega t}\right]
\end{aligned}
$$

Evaluation of the pressure by the method of images was addressed in Chap. 6. The source 1 and its image 2 are depicted in Fig. 2. Because the free surface is pressure-release, the strength of the image is the negative of the source's strength. It is evident that the radial distance $r_{1}$ from the source to the field point equals $\ell_{d}$ and that $r_{2}=\ell_{1}+\ell_{2}=\ell_{r}$.


Figure 2.

The complex pressure amplitude radiated by a point source in a homogenous fluid is $\left[\rho_{0} c \mathcal{P} /(2 \pi)\right]^{1 / 2} \exp (-i k r) / r$, where $\mathcal{P}$ is the time-averaged power. The pressure received at the field point from the image is inverted, so the combination of the source and image pressures is

$$
p(\bar{\xi}, t)=\left(\frac{\rho_{0} c \mathcal{P}}{4 \pi}\right)^{1 / 2} \operatorname{Re}\left[\left(\frac{e^{-i k r_{1}}}{r_{1}}-\frac{e^{-i k r_{2}}}{r_{2}}\right) e^{i \omega t}\right]
$$

Because $r_{1}=\ell_{d}$ and $r_{2}=\ell_{r}$, this is the same result as that obtained from ray theory. Although the analysis using the method of images is much easier, seeing how ray theory is implemented, particularly the analytical evaluation of ray tube area, makes some of the abstract concepts more understandable.

### 11.4 Fermat's Principle

The developments thus far rest on Snell's law, which ultimately is founded on the basic principles that underlie other acoustic phenomena. A different formulation for analyzing ray paths is available. Fermat stated his principle in 1662 without a mathematical derivation. It postulates that the ray connecting two points is such that the path minimizes the travel time between the points. We will see that it leads to Snell's law, as well as a differential equation that is equivalent to the ray tracing equations. However, that is not the reason we study it. Rather it offers a different perspective regarding the nature of rays.

It is possible to justify Fermat's principle with a heuristic argument based on an analogy. This might be the thought process that led Fermat to state the principle. Suppose a large number of people have gathered at the base of a mountain. Each individual is made to follow a different path as they travel to a specified point at the other side of the mountain. Some paths might be easier to walk, so individuals following such paths can travel more quickly. However, these fast paths might be long. Other paths might be much shorter, thereby counterbalancing the fact that those paths are difficult to traverse. If we consider when individuals arrive at the destination, the first arrival will be the individual who followed the optimum path. Individuals who follow paths that are very close to the optimum will arrive slightly later, whereas individuals who follow nonoptimum paths will arrive at a wide range of later times. In the acoustical analogy, signals that follow the optimum and near optimum rays will arrive nearly synchronously, thereby reinforcing each other. Signals following nonoptimal paths will be unsynchronized, and therefore destructively interfere with each other.

Any derivation of basic principles must start from a set of postulates. We could begin with the ray tracing equations as the foundation for a derivation of Fermat's principle. Instead, we will derive the ray tracing equations from it. One benefit of proceeding in this manner is that it will introduce a general mathematical tool called the calculus of variations, which is very useful for a wide range of subjects, especially classical and relativistic mechanics.

In Fig. 11.20 several paths connect selected points $A$ and $B$. One of these paths is optimal in the sense that the time for a signal to travel along it is minimal. Some other paths are far from optimal, but one path in the Figure is very close. This is a variational path. The general approach is to compare the travel time along the optimal path to the time along a variational path. The variational path is defined to be separated from the optimal path by an infinitesimal difference $\delta \bar{\xi}$ of their positions. More specifically, if $\bar{x}=\bar{\xi}$ is a point on the optimal path, then the corresponding point on the variational path is placed at $\bar{x}=\xi+\delta \xi$. The reason for taking $\delta \bar{\xi}$ to be infinitesimal is that doing so allows us to truncate a Taylor series at the second term.

Although we do not know the optimal path, that is, the eigenray, let us pretend that we do. The minimum travel time along this ray between points $A$ and $B$ may be found by evaluating

Fig. 11.20 The optimal ray path corresponds to the minimum time between two designated points $A$ and $B$


$$
\begin{equation*}
T_{\min }=\oint_{A}^{B} \frac{d \ell}{c(\bar{\xi})} \tag{11.4.1}
\end{equation*}
$$

The differential arclength $d \ell$ is the distance between two adjacent points on the optimal path. We shall use the symbol " $d$ " to indicate differential increments pertaining to the same path. Thus, the positions of the adjacent points are $\bar{\xi}$ and $\bar{\xi}+d \bar{\xi}$, as shown in Fig. 11.21. It follows that

$$
\begin{equation*}
d \ell=(d \bar{\xi} \cdot d \bar{\xi})^{1 / 2} \tag{11.4.2}
\end{equation*}
$$

The variational path is obtained by shifting from position $\bar{x}=\bar{\xi}$ to position $\bar{x}=$ $\bar{\xi}+\delta \bar{\xi}$. This corresponds to infinitesimal amount $\delta \bar{\xi}$. The symbol $\delta$ is used to indicate that the increment is infinitesimal, but it pertains to a difference between a quantity on the optimal and variational paths. Thus, any quantity that contains the symbol $\delta$ is not the result of a physical process. For this reason, it is said to be a virtual increment. Thus, $\delta \bar{\xi}$ is a virtual displacement, whereas $d \bar{\xi}$ is an actual displacement. Other than being required to be infinitesimal, the direction and magnitude of the virtual displacement are arbitrary. As shown in Fig. 11.21, $d \bar{\xi}$ is tangent to the optimal path, whereas $\delta \bar{\xi}$ may be in any direction. Because $\delta \bar{\xi}$ is arbitrary, there is no implication that the displaced point, whose position is $\bar{x}=\bar{\xi}+\delta \bar{\xi}$, is on the same wavefront as $\bar{x}=\bar{\xi}$.

Fig. 11.21 Relation between a displacement $d \bar{\xi}$ along minimum-time path and a variational displacement $\delta \bar{\xi}$ to a neighboring sub-optimal path


We need to compare the propagation time in Eq. (11.4.1) to the time for the variational path. This requires a description of the differential arclength along the variational path, which in turn requires that we describe the position of the displaced point on the variational path. Figure 11.21 shows that this position may be regarded as the result of a virtual displacement from the point $\bar{\xi}+d \bar{\xi}$ on the optimal path, or the result of an actual displacement from point $\bar{\xi}+\delta \bar{\xi}$ on the variational path. Both constructions must lead to the same point. An equivalent statement is that the same new position must be obtained, regardless of whether the virtual displacement precedes or follows the actual displacement. Thus, it must be that

$$
\begin{equation*}
(\bar{\xi}+d \bar{\xi})+\delta(\bar{\xi}+d \bar{\xi})=(\bar{\xi}+\delta \bar{\xi})+d(\bar{\xi}+\delta \bar{\xi}) \Longrightarrow \delta(d \bar{\xi})=d(\delta \bar{\xi}) \tag{11.4.3}
\end{equation*}
$$

This property will soon be found to be quite useful.
The arclength along the variational path is the magnitude of the differential increase in the position of a point on the variational path, that is, $d \ell^{\prime}=|d(\bar{\xi}+\delta \bar{\xi})|$. A dot product may be used to construct this distance. Higher order differentials are negligible, so we have

$$
\begin{align*}
d \ell^{\prime} & =[(d \bar{\xi}+\delta(d \bar{\xi})) \cdot(d \bar{\xi}+\delta(d \bar{\xi}))]^{1 / 2} \\
& =[d \bar{\xi} \cdot d \bar{\xi}+2 d \bar{\xi} \cdot \delta(d \bar{\xi})]^{1 / 2} \equiv\left[(d \ell)^{2}+2 d \bar{\xi} \cdot \delta(d \bar{\xi})\right]^{1 / 2}  \tag{11.4.4}\\
& =d \ell+\frac{d \bar{\xi}}{d \ell} \cdot \delta(d \bar{\xi})
\end{align*}
$$

We use this expression to construct the propagation time along the variational path. The sound speed along the variational path is $c(\bar{\xi}+\delta \bar{\xi})$, so the time required for a signal to travel along this path is

$$
\begin{equation*}
T_{\mathrm{var}}=\oint_{A}^{B} \frac{1}{c(\bar{\xi}+\delta \bar{\xi})}\left[d \ell+\frac{d \bar{\xi}}{d \ell} \cdot \delta(d \bar{\xi})\right] \tag{11.4.5}
\end{equation*}
$$

It is easier to proceed if we replace the sound speed with its reciprocal

$$
\begin{equation*}
U(\bar{\xi}) \equiv \frac{1}{c(\bar{\xi})} \tag{11.4.6}
\end{equation*}
$$

(Some treatments use the index of refraction $\mu$ in Eq. (11.2.2), but that would require the unnecessary introduction of a reference sound speed.)

At this juncture, we call on Fermat's principle. According to it, the travel time along the variational path can differ at the most from the minimal time by secondorder differentials. Therefore, we express $T_{\text {var }}$ in terms of the lowest order differentials that distinguish it from $T_{\min }$ in Eq.(11.4.1). A Taylor series is used to treat the integrand, so that

$$
\begin{align*}
T_{\mathrm{var}} & =\oint_{A}^{B} U(\bar{\xi}+\delta \bar{\xi})\left[d \ell+\frac{d \bar{\xi}}{d \ell} \cdot \delta(d \bar{\xi})\right] \\
& =\oint_{A}^{B}[U(\bar{\xi})+\delta \bar{\xi} \cdot \nabla U(\bar{\xi})]\left[d \ell+\frac{d \bar{\xi}}{d \ell} \cdot \delta(d \bar{\xi})\right]  \tag{11.4.7}\\
& =\oint_{A}^{B} U(\bar{\xi}) d \ell+\oint_{A}^{B}\left[U(\bar{\xi}) \frac{d \bar{\xi}}{d \ell} \cdot \delta(d \bar{\xi})+\delta \bar{\xi} \cdot \nabla U(\bar{\xi}) d \ell\right]
\end{align*}
$$

The first integral is $T_{\text {min }}$. Thus, the statement that the lowest order representation of $T_{\text {var }}$ must be the same as $T_{\text {min }}$ leads to the condition that

$$
\begin{equation*}
\oint_{A}^{B}\left[U(\bar{\xi}) \frac{d \bar{\xi}}{d \ell} \cdot \delta(d \bar{\xi})+\delta \bar{\xi} \cdot \nabla U(\bar{\xi}) d \ell\right]=0 \tag{11.4.8}
\end{equation*}
$$

The ray tracing differential equations consider the position of a point on a ray to be a function of time, but none of the terms in the integral contains time. This suggests an alternative description of position on a ray path, in which it is defined by the arclength $\ell$ to a point. In other words, we consider $\ell$ to be the independent variable, so that the position of a point on the optimal path is described functionally as $\bar{x}=\bar{\xi}(\ell)$. This is where Eq. (11.4.3) enters, because we can assert that

$$
\begin{equation*}
\delta(d \bar{\xi})=d(\delta \bar{\xi})=\left[\frac{d}{d \ell}(\delta \bar{\xi})\right] d \ell \tag{11.4.9}
\end{equation*}
$$

This step converts the statement of Fermat's principle in Eq.(11.4.8) to an integral over $\ell$, specifically,

$$
\begin{equation*}
\oint_{A}^{B}\left[U(\bar{\xi}) \frac{d \bar{\xi}}{d \ell} \cdot \frac{d}{d \ell}(\delta \bar{\xi})+\delta \bar{\xi} \cdot \nabla U\right] d \ell=0 \tag{11.4.10}
\end{equation*}
$$

The last operation is based on the fact that although the dependence of $\delta \bar{\xi}$ on $\ell$ is arbitrary, it is a function that we select. Its selection sets $d(\delta \bar{\xi}) / d \ell$, so this derivative is not an independent quantity. It may be eliminated with an integration by parts, which gives

$$
\begin{equation*}
\left.U(\bar{\xi}) \frac{d \bar{\xi}}{d \ell} \cdot(\delta \bar{\xi})\right|_{\ell_{A}} ^{\ell_{B}}-\oint_{A}^{B}\left[\frac{d}{d \ell}\left(U(\bar{\xi}) \frac{d \bar{\xi}}{d \ell}\right)-\nabla U\right] \cdot(\delta \bar{\xi}) d \ell=0 \tag{11.4.11}
\end{equation*}
$$

At this juncture, we take an overview. It was stated that $\delta \bar{\xi}$ is an arbitrary function of $\ell$, but that is not exactly true. The variational path is an alternative between point $A$ and $B$, so the variational path must intersect these points. This means that $\delta \bar{\xi}\left(\ell_{A}\right)=\delta \bar{\xi}\left(\ell_{B}\right)=0$. The consequence is that the terms in the preceding relation that are evaluated at the ends of the path are zero. In turn, this requires that the integral must evaluate to zero for any $\delta \bar{\xi}$ function we select. If we wish that the integral of
$\bar{a} \cdot \bar{b}$ be zero for any function $\bar{b}$, it must be that $\bar{a}$ is identically zero. Thus, we find

$$
\begin{equation*}
\frac{d}{d \ell}\left(U(\bar{\xi}) \frac{d \bar{\xi}}{d \ell}\right)-\nabla U=0 \tag{11.4.12}
\end{equation*}
$$

Minimization of integral quantities arise in a variety of areas, notably analytical dynamics. The mathematical process we have followed is known as the calculus of variations. Equation (11.4.12) in a general analysis is known as the EulerLagrange equation. It is a second-order differential equation governing $\bar{\xi}$, which is a fundamental difference with the (first order) ray tracing equations. Another fundamental difference is the absence of the wave slowness $\bar{s}$.

Our intent at the outset was to derive the ray tracing equations from Fermat's principle, but the appearance of Eq. (11.4.12) does not suggest how to meet that objective. Let us try to work backward, by eliminating $\bar{s}$ from the ray tracing equations. The second ray tracing equation, Eq. (11.3.10), states that

$$
\begin{equation*}
\frac{d \bar{\xi}}{d t}=c^{2} \bar{s} \tag{11.4.13}
\end{equation*}
$$

A different view is that $\bar{\xi}$ depends on the arclength $\ell$, which leads to a description of the preceding derivative according to $d \bar{\xi} / d t=(d \bar{\xi} / d \ell)(d \ell / d t)$. The local sound speed is the rate at which the arclength changes, so that

$$
\begin{equation*}
c=\frac{d \ell}{d t} \tag{11.4.14}
\end{equation*}
$$

Matching the two descriptions of the rate of change of the position on a ray leads to

$$
\begin{equation*}
\frac{d \bar{\xi}}{d t}=\frac{d \bar{\xi}}{d \ell} \frac{d \ell}{d t} \equiv \frac{d \bar{\xi}}{d \ell} c=c^{2} \bar{s} \tag{11.4.15}
\end{equation*}
$$

Replacing $c$ with $1 / U$ thereby converts the second ray tracing equation to

$$
\begin{equation*}
\bar{s}=U \frac{d \bar{\xi}}{d \ell} \tag{11.4.16}
\end{equation*}
$$

Now we turn to the first ray tracing equation, Eq. (11.3.9), which is

$$
\begin{equation*}
\frac{d \bar{s}}{d t}=-\frac{\nabla c}{c} \tag{11.4.17}
\end{equation*}
$$

On the right side, we set $c=1 / U$, so that

$$
\begin{equation*}
\frac{\nabla c}{c}=U \nabla\left(\frac{1}{U}\right)=-\frac{1}{U} \nabla U \tag{11.4.18}
\end{equation*}
$$

We introduce the chain rule to handle $d \bar{s} / d t$ on the left side of Eq.(11.4.17) because our view now is that $\bar{s}$ is a function of $\ell$. The result is that the first ray tracing equation becomes

$$
\begin{equation*}
\frac{d \bar{s}}{d \ell} \frac{d \ell}{d t} \equiv \frac{d \bar{s}}{d \ell} c=\frac{1}{U} \nabla U \tag{11.4.19}
\end{equation*}
$$

Replacement of $c$ with $1 / U$ in this equation leads to

$$
\begin{equation*}
\frac{d \bar{s}}{d \ell}=\nabla U \tag{11.4.20}
\end{equation*}
$$

The last step is to substitute $\bar{s}$ from Eq. (11.4.16) into this expression. The result is the Euler-Lagrange equation, Eq. (11.4.12). In other words, the ray tracing equations and the Euler-Lagrange equation are equivalent specifications of a ray path. The differences are that the ray tracing equations give position as a function of time, with additional information about the tangent to the path in the form of $\bar{s}$. If we wish to also determine the arclength to a point on a ray, we can do so by adding $\ell$ to the set of unknowns $\{X\}$ in Eq. (11.3.12) and inserting into the scalar differential equations another equation: $d \ell / d t=c\left(\xi_{x}, \xi_{y}, \xi_{z}\right)$. The Euler-Lagrange equation describes the position in terms of $\ell$. If we wish to know a local tangent or the propagation time, we must extract that information after the equation has been solved.

Some mathematical analyses are best performed with a single differential equation. However, realistic situations usually require numerical methods. In that case, the ray tracing equations are preferable. The reason for this assertion is that most numerical algorithms require that the equations be first order, which is how the ray tracing equations are posed. Of course, the Euler-Lagrange equation could be transformed into first-order form, but as we have seen, such equations would be the ray tracing equations or a variant of them.

We have demonstrated that Fermat's principle correctly describes the rays in a fluid whose sound speed is a continuous function of position. It also is valid at discontinuities. To demonstrate this, consider Fig. 11.22. We wish to identify the horizontal distance $x_{O}$ where the ray from $A$ to $B$ crosses the interface between two fluids. Points $A$ and $B$ are fixed. A coordinate system whose $z$-axis measures depth into the second fluid is situated at the horizontal position of $A$.

Fig. 11.22 Construction of a ray path for transmission across an interface between two fluids


The travel times along each segment of the ray are $\ell_{1} / c_{1}$ and $\ell_{2} / c_{2}$. These distances are

$$
\begin{equation*}
\ell_{1}=\left[z_{A}^{2}+\left(x_{O}\right)^{2}\right]^{1 / 2}, \quad \ell_{2}=\left[z_{B}^{2}+\left(x_{B}-x_{O}\right)^{2}\right]^{1 / 2} \tag{11.4.21}
\end{equation*}
$$

We seek the value of $x_{O}$ that minimizes the sum of the travel time along each segment. This requires that

$$
\begin{align*}
\frac{d}{d x_{O}}\left(\frac{\ell_{1}}{c_{1}}+\frac{\ell_{2}}{c_{2}}\right) & =\frac{d}{d x_{O}} \frac{\left[z_{A}^{2}+\left(x_{O}\right)^{2}\right]^{1 / 2}}{c_{1}}+\frac{d}{d x_{O}} \frac{\left[z_{B}^{2}+\left(x_{B}-x_{O}\right)^{2}\right]^{1 / 2}}{c_{2}}  \tag{11.4.22}\\
& =\frac{x_{O}}{c_{1} \ell_{1}}-\frac{\left(x_{B}-x_{O}\right)}{c_{2} \ell_{2}}=0
\end{align*}
$$

According to Fig. 11.22, $\sin \psi_{1}=x_{O} / \ell_{1}$ and $\sin \psi_{2}=\left(x_{B}-x_{0}\right) / \ell_{2}$, so the preceding is a restatement of Snell's law.

There is a degree of elegance in having a general principle that leads to the equations for ray paths in inhomogeneous fluids as well as at interfaces. Indeed, the generality of Fermat's principle probably is its most important attribute. For example, it can been used to derive equations describing the rays in the presence of an ambient flow. ${ }^{11}$ It also can link geometrical acoustics to optics and the Hamiltonian formulation of analytical mechanics.

EXAMPLE 11.7 A point source is located at position $A$ in an ocean channel and a hydrophone is situated at point $B$. The sound speeds are $c_{1}$ for the water, $c_{2}<c_{1}$ for the air, and $c_{3}>c_{1}$ for the sediment. Describe the eigenrays for point $B$.


Figure 1.

## Significance

This application of Fermat's principle to the determination of eigenrays is quite straightforward, but the results have a surprising feature. Thus, this example affirms that the perspective offered by Fermat's principle might cause us to re-examine our initial assumptions.

[^55]
## Solution

We begin with a sketch of possible paths between points $A$ and $B$. Paths 1,2 , and 3 in Fig. 2 are familiar, but path 4 might be considered to be unexpected. In addition to these, others exist, but we will consider them later. Path 1 is the direct path. Because the sound speed along it is $c_{1}$, the minimum travel time is obtained for the shortest path. Thus, Fermat's principle confirms that this path is a straight line.


Figure 2.

Paths 2 and 3 feature a single reflection. Let us determine whether Fermat's principle yields ray paths for which the angle of reflection equals the angle of incidence. In the case of the surface-reflected wave, points $A$ and $B$ are fixed and we know that the elevation of point $C$ at which it is incident. This ray path is set if the horizontal location of point $C$ is specified. Therefore, we select the horizontal range $R_{C}$ to this point as the unknown to be determined by Fermat's principle. Trigonometric equations describe the length of the two straight segments that form path 2. Division of those lengths by the sound speed gives the travel time along each. Fermat's principle applies to the sum of these contributions, which is

$$
T_{2}\left(R_{C}\right)=\frac{\left(d_{A}^{2}+R_{C}^{2}\right)^{1 / 2}}{c_{1}}+\frac{\left[d_{B}^{2}+\left(R-R_{C}\right)^{2}\right]^{1 / 2}}{c_{1}}
$$

Fermat's principle states that the correct value of $R_{C}$ is that which gives $d T_{2} / d R_{C}=0$. Evaluation of the derivative gives

$$
\begin{equation*}
\frac{d T_{2}}{d R_{C}}=\frac{1}{c_{1}} \frac{R_{c}}{\left(d_{A}^{2}+R_{C}^{2}\right)^{1 / 2}}-\frac{1}{c_{1}} \frac{\left(R-R_{C}\right)}{\left[d_{B}^{2}+\left(R-R_{C}\right)^{2}\right]^{1 / 2}}=0 \tag{1}
\end{equation*}
$$

The length ratios in Eq. (1) are the cosines of the respective grazing angle depicted in the sketch. Hence, Eq. (1) reduces to

$$
\frac{\cos \chi_{2, A}}{c_{1}}-\frac{\cos \chi_{2, B}}{c_{1}}=0
$$

Cancelation of $c_{1}$ shows that the grazing angles are equal. Thus, Fermat's principle is consistent with the familiar reflection law. There is no need to repeat this analysis for path 3, because the arrangement is the same as path 2 .

Path 4 involves two fluids. It consists of a straight segment from the source to point $E$ on the interface, then a horizontal path in the sediment to point $F$, from which a straight path brings the signal to point $B$. The location of points $E$ and $F$ each depend on the horizontal range, so this path is a function of two parameters. For the sake of variety, let us use the grazing angles $\chi_{A, 4}$ and $\chi_{B, 4}$. The corresponding expression for the travel time is

$$
T_{4}\left(\chi_{A, 4}, \chi_{B, 4}\right)=\frac{1}{c_{1}} \frac{H-d_{A}}{\sin \chi_{A, 4}}+\frac{1}{c_{3}}\left[R-\left(\frac{H-d_{A}}{\tan \chi_{A, 4}}\right)-\left(\frac{H-d_{B}}{\tan \chi_{B, 4}}\right)\right]+\frac{1}{c_{1}} \frac{H-d_{B}}{\sin \chi_{B, 4}}
$$

The value of $T_{4}$ must be stationary with respect to both grazing angles, so we set

$$
\begin{align*}
\frac{\partial T_{4}}{\partial \chi_{A, 4}} & =\left(\frac{H-d_{A}}{c_{1}}\right) \frac{\left(-\cos \chi_{A, 4}\right)}{\left(\sin \chi_{A, 4}\right)^{2}}-\left(\frac{H-d_{A}}{c_{3}}\right) \frac{(-1)}{\left(\sin \chi_{A, 4}\right)^{2}}=0  \tag{2}\\
\frac{\partial T_{4}}{\partial \chi_{B, 4}} & =-\left(\frac{H-d_{B}}{c_{2}}\right) \frac{(-1)}{\left(\sin \chi_{B, 4}\right)^{2}}+\left(\frac{H-d_{B}}{c_{3}}\right) \frac{\left(-\cos \chi_{B, 4}\right)}{\left(\sin \chi_{B, 4}\right)^{2}}=0
\end{align*}
$$

These expressions reduce to

$$
\cos \chi_{A, 4}=\frac{c_{1}}{c_{3}}=\cos \chi_{B, 4}
$$

The angles of the line segments relative to the normal to the interface are $\psi_{A, 4}=$ $\pi / 2-\chi_{A, 4}$ and $\psi_{B, 4}=\pi / 2-\chi_{B, 4}$, so Eq. (2) leads to

$$
\begin{equation*}
\sin \psi_{A, 4}=\sin \psi_{B, 4}=\frac{c_{1}}{c_{3}} \tag{3}
\end{equation*}
$$

This is the expression for the critical angle of incidence. It is a real value only if the sound speed in the receiving fluid is greater than the speed in the transmitting fluid. Thus, Fermat's principle tells us that the alternative path exists if $c_{3}$ for the sediment is greater than $c_{1}$ for the water. If that is the case, then the ray arrives and departs from the interface at the critical angle. One way to view this is that the source at point $A$ emits many rays. Rays that arrive at the sediment at an angle of incidence less than the critical value are reflected upward and transmitted downward. However, the ray that arrives at the critical angle of incidence generates a transmitted ray that is parallel to the interface. This ray runs along the interface on the sediment side. The water now becomes the receiving fluid. The ray emerges from the sediment when the angle from the interface to the receiving point is the critical angle for transmission from the fluid into the sediment.

Paths 2 and 3 entail a single reflection. Multiple reflection paths also are possible. A few are shown in Fig.3. In general, there the two paths that feature $N$ reflections, with the their difference being whether the first reflection is at the top or bottom. We will analyze ray 5 , from which the procedure for other cases should be apparent.


Figure 3.

The reflection points are defined by the ranges $R_{C}$ and $R_{D}$, so the travel time is a function of two parameters,

$$
\begin{aligned}
T_{5}\left(R_{C}, R_{D}\right)= & \frac{\left[\left(d_{A}\right)^{2}+\left(R_{C}\right)^{2}\right]^{1 / 2}}{c_{1}}+\frac{\left[H^{2}+\left(R_{D}-R_{C}\right)^{2}\right]^{1 / 2}}{c_{1}} \\
& +\frac{\left[\left(H-d_{B}\right)^{2}+\left(R-R_{D}\right)^{2}\right]^{1 / 2}}{c_{1}}
\end{aligned}
$$

For an extreme value, the derivative of $T_{5}$ with respect to both parameters must be zero,

$$
\begin{gather*}
\frac{\partial T_{5}}{\partial R_{C}}=\frac{1}{c_{1}} \frac{R_{C}}{\left[\left(d_{A}\right)^{2}+\left(R_{C}\right)^{2}\right]^{1 / 2}}-\frac{1}{c_{1}} \frac{\left(R_{D}-R_{C}\right)}{\left[H^{2}+\left(R_{D}-R_{C}\right)^{2}\right]^{1 / 2}}=0 \\
\frac{\partial T_{5}}{\partial R_{D}}=\frac{1}{c_{1}} \frac{\left(R_{D}-R_{C}\right)}{\left[H^{2}+\left(R_{D}-R_{C}\right)^{2}\right]^{1 / 2}}-\frac{1}{c_{1}} \frac{\left(R-R_{D}\right)}{\left[\left(H-d_{B}\right)+\left(R-R_{D}\right)^{2}\right]^{1 / 2}}=0 \tag{4}
\end{gather*}
$$

These equations may be solved for $R_{C}$ and $R_{D}$. To interpret these equations we observe that the grazing angle of the segments of ray 5 in the Figure are

$$
\begin{gather*}
\chi_{A C}=\cos ^{-1}\left(\frac{R_{C}}{\left[\left(d_{A}\right)^{2}+\left(R_{C}\right)^{2}\right]^{1 / 2}}\right) \\
\chi_{C D}=\cos ^{-1}\left(\frac{R_{D}-R_{C}}{\left[H^{2}+\left(R_{D}-R_{C}\right)^{2}\right]^{1 / 2}}\right)  \tag{11.4.23}\\
\chi_{A C}=\cos ^{-1}\left(\frac{R-R_{D}}{\left[\left(H-d_{B}\right)^{2}+\left(R-R_{D}\right)^{2}\right]^{1 / 2}}\right)
\end{gather*}
$$

Thus, the values of $R_{C}$ and $R_{D}$ that extremize $T_{5}$ according to Eq. (4) correspond to equal grazing angles for the arriving and departing rays at points $C$ and $D$. In other words, the angle of reflection equals the angle of incidence at points $C$ and $D$. There is nothing new in this.

The fact that Fermat's principle identifies the critical angle path through the sediment is another demonstration of its generality. One of the paradoxes of this path is
that the concept of a ray tube having an identifiable cross-sectional area breaks down. This is so because grazing angles less than the critical value in Eq. (3) do not lead to real rays, while grazing angles greater than critical lead to rays that are transmitted into the sediment, and therefore cannot reach the vicinity of point $B$. This is one of many phenomena that do not fit the geometric acoustic model. A few others are the fact that a sound is heard in a shadow zone, and that sound is heard behind a rigid barrier. Phenomena that do not fit into the geometrical acoustic model are referred to as diffraction. Analyses of diffraction phenomena generally require a level of mathematics that is beyond the scope of this book.

### 11.5 Homework Exercises

Exercise 11.1 The mirror in the sketch is a surface of revolution $z=f(R)$, with $f(R)$ increasing monotonically. A plane wave propagating in the negative $z$ direction is reflected. Consider a ray that is incident at arbitrary $R$. The angle of reflection equals the angle of incidence, so $\tan \psi=d f / d R$. Derive an expression in terms of $R, f(r)$, and $d f / d R$ for the coordinate $z_{f}$ where the reflection of this ray intersects the $z$-axis. Then consider the specific case of a paraboloid of revolution, for which $f(R)=h_{z}(R / a)^{2}$. Show that $z_{f}$ is a constant in this case, which means that all reflected rays intersect at a true focus.


Exercise 11.1

Exercise 11.2 The cross section of the prism in the sketch is an isosceles triangle with a vertical axis of symmetry. Assume that the prism is a fluid with sound speed $c_{b}$, and set $c_{a}$ as the sound speed in the surrounding fluid. (a) Derive an expression for the angle of transmission $\psi$ for rays that emerge from the prism. (b) If $c_{a}=200 \mathrm{~m} / \mathrm{s}$ and $c_{b}=400 \mathrm{~m} / \mathrm{s}$, what is the largest value of $\phi$ for which a plane wave emerges from the prism? (b) If $c_{a}=400 \mathrm{~m} / \mathrm{s}$ and $c_{b}=200 \mathrm{~m} / \mathrm{s}$, what is the largest value of $\phi$ for which a plane wave emerges from the prism? (c) It is given that $c_{a} \neq c_{b}$. Qualitatively explain why there is no value of the apex angle $2 \phi>0$ for which the plane wave that emerges propagates horizontally, regardless of whether $c_{a}>c_{b}$ or $c_{a}<c_{b}$.


Exercise 11.2

Exercise 11.3 The reflection from a flat surface of the signal radiated by a point source is describable by the method of images if the surface is rigid or pressurerelease. In the case of a locally reacting surface, geometrical acoustics also leads to the conclusion that the reflected rays seem to emanate from a point source below the surface. It also leads to the conclusion that the image source has a directivity factor. Consider an omnidirectional point source situated at height $H$ above a surface whose local impedance is $Z$. The source is harmonic at frequency $\omega$. Determine the location of the image and the directivity of that image. Does the field of the image source actually satisfy the Helmholtz equation in the actual fluid?

Exercise 11.4 The sound speed profile in a very deep ocean channel consists of a lower layer of depth $H$ in which the sound speed is constant at $c_{1}$, and an upper layer in which the sound speed decreases linearly with depth. The sound speed is continuous across the interface of these layers. It is convenient to describe $c$ in terms of the height $z^{\prime}$ above the bottom. Thus, the sound speed is $c=c_{1}+m\left(z^{\prime}-H\right) h\left(z^{\prime}-H\right)$, with $m>0$ and $z^{\prime}=0$ at the bottom. Any ray path within the upper layer must be a downward curving circular arc. Consequently, a ray departing from the bottom at any upward grazing angle must eventually return to the bottom. Derive an expression for the distance between the points of departure and return.

Exercise 11.5 At the surface of an ocean channel, the sound speed is $1470 \mathrm{~m} / \mathrm{s}$. It decreases linearly with depth, down to 500 m , where the sound speed is $1420 \mathrm{~m} / \mathrm{s}$. Downward from that depth the sound speed increases linearly, with a value of $1470 \mathrm{~m} / \mathrm{s}$ occurring again at 1000 m . Determine the path of a ray launched horizontally at a depth of 200 m . Base the analysis on the fact the path is circular if the sound speed is linearly dependent on the depth.

Exercise 11.6 External heating of a waveguide has resulted in a sound speed that increases linearly from axial position $x=0$ to position $x=L$, but the sound speed is constant transversely. The result is that $c=c_{0}(1-x / L)+c_{1} x / L$. A point source is situated in the middle of the waveguide, at $x=0, y=H / 2$, where $H=2 \mathrm{~m}$ is the waveguide's width. Consider rays launched at grazing angle of $15^{\circ}, 45^{\circ}$, and $75^{\circ}$. (Note that the heterogeneity varies in the $x$ direction, so the grazing angle are measured relative to the $y$ direction.) Do any of these rays intersect a wall? If so,
what is the location of this intersection, and what is the propagation time from the instant the ray is launched to the instant when it arrives at that wall? Parameters are $c_{0}=300 \mathrm{~m} / \mathrm{s}, c_{1}=400 \mathrm{~m} / \mathrm{s}, H=3 \mathrm{~m}$.

Exercise 11.7 The sound speed in a hard-walled two-dimensional HVAC duct varies parabolically according to $c(z)=c_{0}+c_{1}(1-2 z / H)^{2}$, where $z$ is the transverse distance. The walls are at $y=0$ and $y=H$. Determine the path of two rays that are launched downward from $z=H / 2: \chi_{\mathrm{Ia}}$ is such that a wall is a turning point, and $\chi_{\mathrm{Ib}}=0.5 \chi_{\mathrm{Ia}}$. Parameters are $c_{0}=300 \mathrm{~m} / \mathrm{s}, c_{1}=50 \mathrm{~m} / \mathrm{s}, H=2 \mathrm{~km}$.

Exercise 11.8 The sound speed in a hard-walled two-dimensional HVAC duct varies parabolically according to $c(z)=c_{0}+c_{1}(1-2 z / H)^{2}$, where $z$ is the transverse distance. The walls are at $z=0$ and $z=H$. Determine the paths of rays launched downward from $z=H / 2$ resulting from an initial grazing angle that is $1.25 \chi_{\text {Iw }}$, where $\chi_{\mathrm{Iw}}$ is such that a wall is a turning point. Parameters are $c_{0}=300 \mathrm{~m} / \mathrm{s}, c_{1}=50 \mathrm{~m} / \mathrm{s}$, $H=2 \mathrm{~km}$.

Exercise 11.9 The sound speed in a large body of water varies sinusoidally with depth relative to a mean value, $c(z)=c_{0}+\varepsilon \sin (2 \pi z / H)$. A radially symmetric point source is situated at $z_{0}=10 \mathrm{H}$. This depth is sufficiently large that there is no need to consider reflections from the surface or bottom. (a) Consider a ray that is launched horizontally. Does it have a turning point? If so, at what depth? (b) Determine the path followed by this ray. (c) Determine the travel time along this ray. Plot this property as a function of the horizontal distance from the source, and also as a function of the depth. Parameters are $c_{0}=1500 \mathrm{~m} / \mathrm{s}, \varepsilon=25 \mathrm{~m} / \mathrm{s}, H=100 \mathrm{~m}$.

Exercise 11.10 The sound speed in a large body of water varies sinusoidally with depth relative to a mean value, $c(z)=c_{0}+\varepsilon \sin (2 \pi z / H)$. A radially symmetric point source is situated at $z_{0}=10 \mathrm{H}$. This depth is sufficiently large that there is no need to consider reflections from the surface or bottom. Determine the paths of rays launched at initial grazing angles ranging from $-45^{\circ}$ to $45^{\circ}$ in a $1^{\circ}$ increment.

Exercise 11.11 Consider the waveguide in Exercise 11.10. (a) Evaluate the paths of rays that are launched from the source at grazing angles of $2.99^{\circ}$ and $3.01^{\circ}$. (b) Use the result of Part (a) to evaluate the ray tube area corresponding to rays that depart from in the source range $2.99^{\circ} \leq \chi \leq 3.01^{\circ}$ in an interval of azimuthal angles subtending $0.02^{\circ}$. (c) Determine the intensity in this ray tube at a horizontal distance of 500 m from the source.

Exercise 11.12 Evaluate the ray tube area $\mathcal{A}$ for two incident rays in Exercise 11.2 that are infinitesimally separated by elevation $d h$. From that expression determine the amplitude $\left|P_{\mathrm{T}}\right|$ of the wave that emerges from the prism if the incident wave was harmonic with amplitude $\left|P_{\mathrm{I}}\right|$. Graph $d \mathcal{A} / d h$ and $\left|P_{\mathrm{T}}\right| /\left|P_{\mathrm{I}}\right|$ as functions of $\theta$ for two cases: $c_{b} / c_{a}=2, \rho_{b} / \rho_{a}=1.2$, and $c_{b} / c_{a}=0.5, \rho_{b} / \rho_{a}=0.8$.

Exercise 11.13 A radially symmetric source is situated at depth $z_{0}$ in a vertically stratified fluid. The sound speed is an arbitrary function $c(z)$, where $z$ is the depth.

It is desired to determine the intensity within a ray tube that surrounds the vertical ray path. The first step in the analysis considers a ray that departs from the source slightly off vertical, so the initial grazing angle is $\chi_{\mathrm{I}}=\pi / 2-\Delta$, with $\Delta \ll 1$.
(a) Show that the slope of this ray at an arbitrary depth is given, to leading order in powers of $\Delta$, by

$$
\frac{d R}{d z} \approx \Delta \frac{c(z)}{c_{\mathrm{I}}}
$$

where $R$ is the transverse distance from the $z$-axis. (b) This approximation is valid at any depth for which $c(z) / c_{\mathrm{I}}$ is $O(1)$. Subject to this limitation derive an expression for the cross section area of the ray tube that emanates from the source as a cone whose apex angle is $2 \Delta$. Then use that area to determine the depth dependence of the time-averaged intensity along the downward ray.

Exercise 11.14 Consider the waveguide in Exercise 11.6. The task here is to determine the pressure at the wall when the radially symmetric source, which is situated at $x=0, y=H / 2$, is time-harmonic, with time-averaged power output $\mathcal{P}$. Derive an expression for the distance $x_{f}$ at which a ray that departs from the source at grazing angle $\chi_{\mathrm{I}}$ intersects a wall. Then use that expression to determine the pressure amplitude at the location where rays intersect the wall. Consider initial grazing angles of $\chi_{\mathrm{I}}=0^{\circ}, 30^{\circ}$, and $60^{\circ}$. (Note that the heterogeneity varies in the $x$ direction, so the grazing angle are measured relative to the $y$ direction.)

Exercise 11.15 Suppose that the reference sound speeds in Example 11.5 are reversed, so that $c_{0}=280 \mathrm{~m} / \mathrm{s}, c_{1}=330 \mathrm{~m} / \mathrm{s}$. All other parameters are as stated there. Compute the ray paths and wavefronts. Do caustics form? Are there shadow zones?

Exercise 11.16 Due to extraordinary circumstances the sound speed in the atmosphere has been found to depend on the horizontal distance $x$ and altitude $y$ measured from a point that is 1000 m above the ground. The dependence $c=350+x y / 200 \mathrm{~m} / \mathrm{s}$, where the units of $x$ and $y$ are meters. A weather balloon transmits an omnidirectional point source from that location. (a) Compute the paths of rays that are launched in the xy plane in all directions covering the full $360^{\circ}$ range in $10^{\circ}$ increments. (b) Determine the wavefronts corresponding to the rays in Part (a).

Exercise 11.17 In the region $x<0$, a fluid is homogeneous, with sound speed $c_{0}$. To the right, $x>0$, the sound speed varies sinusoidally in both the $x$ and $y$ directions, according to $c=c_{0}-\Delta \sin (\pi x / L) \cos (\pi y / H)$ for $0<x<L$. For $x>L$ the sound speed returns to $c=c_{0}$. The parameters are $c_{0}=1500 \mathrm{~m} / \mathrm{s}, \Delta=60 \mathrm{~m} / \mathrm{s}, L=3 \mathrm{~m}$, $H=1.5 \mathrm{~m}$. An initially plane wave whose propagation direction is $\bar{e}=\bar{e}_{x}$ is incident at $x=0$ on the heterpgeneous region. Determine the ray paths and wavefronts in the region $0>x>6 L$.

Exercise 11.18 Consider the fluid in Exercise 11.17. A harmonic plane wave having amplitude $B$ propagates at $8^{\circ}$ above the $x$-axis in the region $x<0$. Consider the ray that is incident on the heterogeneous region at $x=0, y=H / 4$. (a) Determine the
path of this ray. (b) What is the location of this ray and the direct of the tangent to that ray when it emerges from the heterogeneous region at $x=L$. (c) Determine the amplitude of the signal described by this ray at the location where the ray emerges from the heterogeneous regions

Exercise 11.19 A source at point $A$ emits rays in all directions. Use Fermat's principle to prove that the ray that arrives at field point $B$ after reflection from the corner is parallel to the ray that reflects from the vertical wall.


Exercise 11.19

Exercise 11.20 The sketch shows two points at opposite sides of a rigid cylinder. The fluid in which it is immersed is homogeneous, with constant sound speed $c$. Suppose there is a ray connecting these points. Because path segments within the fluid must be straight, the only way a connecting ray might exist is if it wraps around the surface of the cylinder, as shown in the sketch. Given distances $L_{1}$ and $L_{2}$ use Fermat's principle to determine the angles $\phi_{1}$ and $\phi_{2}$ that mark the limits of this contact. Draw a sketch of the result. Also determine the corresponding propagation time for a signal that follows this path.


Exercise 11.20

## Chapter 12 Scattering

Have you ever wondered why you can hear sound even though there is a building between you and the source, or why the sky is blue, or how dolphins can identify objects underwater? These phenomena are manifestations of scattering. At a fundamental level, scattering is the reflection of a signal from a finite-sized object. Knowledge of the scattering properties of a body probably is of the greatest concern for underwater acoustics, not only for its relevance to naval applications like sonar, but also for animal bioacoustics. Scattering phenomena often also are important for studies of atmospheric propagation.

We will see that sometimes the analysis of scattering is closely related to the description of acoustic radiation. Like that problem, scattering analyses are usually performed in the frequency domain. The key parameter is the wavelength relative to the size of the scattering region. The primary tools for low frequencies are derived from the Kirchhoff-Helmholtz integral theorem, whereas geometrical acoustics is the approach we will employ for high frequencies. Both formulations entail the judicious application of well-justified approximations.

The problematic situation is the mid-frequency range, in which the object's dimensions are comparable to the wavelength. There are few objects that are conducive to an analytical study of the scattered field at an arbitrary frequency. The sphere is one that we shall take up. Spherical harmonics will play an important role in that study. As is true for many simplified models, the availability of this solution will help us understand the validity of the low- and high-frequency approximations. Nevertheless, we seldom encounter spheres as actual acoustical systems. Realistic shapes require computational methods.

### 12.1 Background

The general situation has two finite-sized bodies that are surrounded by a fluid whose extent is infinite. One generates a signal that encounters the other body as an obstacle to its propagation. The modification of the incident signal is the scattered signal. Reflection is one aspect of the scattering process, but other phenomena are involved. In the most accurate model, the pressure distribution on the surface of the scattering body causes that body to move. Such motion modifies the scattered signal in comparison with what it would be if the body were immobile. Incorporation of the effect of that motion might require consideration of the laws of structural dynamics. Another complication is that the scattered signal returns to the generating body. Consequently, that body also is a scatterer, so its surface response is different from that associated with generation of the incident signal. For example, if the generating body is an electromechanical transducer, the pressure and particle velocity corresponding to a specified voltage are modified. This alters the signal that is incident on the second body that was considered at the outset to be the scatterer, and so on. This is the phenomenon of mutual scattering. The extended version features many bodies whose scattered fields interact.

An analysis of mutual scattering is quite complicated and requires prior understanding of single-body scattering. The latter is a very good model if the distance between the two bodies is much greater than the acoustic wavelength and the largest dimension of either body, because the farfield of either scattered signal will eventually decay inversely with the distance from that body. The consequence is that the signal radiated by the generating body will be little affected by the signal returning from the scatterer.

Thus, the general problem we shall address is that body $\mathcal{S}_{0}$ generates the signal $\operatorname{Re}\left(P_{\mathrm{I}}(\bar{x}) e^{i \omega t}\right)$ that is incident on the scattering body $\mathcal{S}$. The scattered signal is $\operatorname{Re}\left(P_{\mathrm{S}}(\bar{x}) e^{i \omega t}\right)$. The total pressure is the sum of these contributions,

$$
\begin{equation*}
P(\bar{x})=P_{\mathrm{I}}(\bar{x})+P_{\mathrm{s}}(\bar{x}) \tag{12.1.1}
\end{equation*}
$$

As noted, because $\mathcal{S}$ and $\mathcal{S}_{0}$ are widely separated, the field generated by $\mathcal{S}_{0}$ is essentially the same as it would be if $\mathcal{S}$ were not present. Thus, $P_{\mathrm{I}}$ is a solution of the Helmholtz equation that matches the normal velocity on $\mathcal{S}_{0}$, and obeys the Sommerfeld radiation condition at large radial distances $r_{0}$ from $\mathcal{S}_{0}$,

$$
\begin{gather*}
\nabla^{2} P_{\mathrm{I}}+k^{2} P_{\mathrm{I}}=0 \\
\bar{n} \cdot \nabla P_{\mathrm{I}}=-i \omega \rho_{0} V_{0}, \quad \bar{x} \in \mathcal{S}_{0} \\
\lim _{r_{0} \rightarrow \infty} r_{0}\left(\frac{\partial P_{\mathrm{I}}}{\partial r_{0}}+i k P_{\mathrm{I}}\right)=0 \tag{12.1.2}
\end{gather*}
$$

In principle, the incident field can be any field consistent with these equations. However, we have imposed the restriction that $\mathcal{S}$ is very distant from $\mathcal{S}_{0}$, which means that it is reasonable in most cases to consider the field that is incident on $\mathcal{S}$ to be locally planar.

If the scattering body is capable of motion, either as a rigid body, or as a consequence of its deformability, then the surface velocity on $\mathcal{S}$ is $\bar{V}(\bar{x})$. The normal velocity in the acoustic signal must match the normal component of $V(\bar{x})$. Also, the total pressure must satisfy the Helmholtz equation. Because $P$ is the superposition in Eq. (12.1.1) and $P_{\mathrm{I}}$ individually satisfies the Helmholtz equation, the scattered pressure must satisfy

$$
\begin{gather*}
\nabla^{2} P_{\mathrm{s}}+k^{2} P_{\mathrm{s}}=0  \tag{12.1.3}\\
\bar{n} \cdot \nabla P_{\mathrm{s}}=-\bar{n} \cdot \nabla P_{\mathrm{I}}-i \omega \rho_{0} \bar{n} \cdot \bar{V}, \quad \bar{x} \in \mathcal{S}
\end{gather*}
$$

In addition, the scattered field at large radial distances $r$ from $\mathcal{S}$ must satisfy the Sommerfeld radiation condition relative to that body,

$$
\begin{equation*}
\lim _{r \rightarrow \infty} r\left(\frac{\partial P_{\mathrm{s}}}{\partial r}+i k P_{\mathrm{s}}\right)=0 \tag{12.1.4}
\end{equation*}
$$

These equations may be interpreted as a statement that the scattered field is the pressure radiated by $\mathcal{S}$ when the normal velocity of its surface is $\bar{n} \cdot \bar{V}+\left(1 / i \omega \rho_{0}\right) \bar{n}$. $\nabla P_{\mathrm{I}}$.

By definition, the scattered pressure is the field created by the insertion of $\mathcal{S}$ into the fluid domain. The pressure on that surface is the total quantity $P(\bar{x})$. By analogy with a radiation problem, the Kirchhoff-Helmholtz integral theorem (KHIT) describes $P_{\mathrm{s}}$ at any location $\bar{x}_{0}$ exterior to $\mathcal{S}$. Equation (7.4.14) states that

$$
\begin{equation*}
P_{\mathrm{s}}\left(\bar{x}_{0}\right)=\iint_{\mathcal{S}}\left[P\left(\bar{x}_{s}\right) \bar{n}\left(\bar{x}_{s}\right) \cdot \nabla G\left(\bar{x}_{0}, \bar{x}_{s}\right)-\bar{n}\left(\bar{x}_{s}\right) \cdot \nabla P\left(\bar{x}_{s}\right) G\left(\bar{x}_{0}, \bar{x}_{s}\right)\right] d \mathcal{S} \tag{12.1.5}
\end{equation*}
$$

Although KHIT is valid for any Green's function, the one we will use is the free-space version,

$$
\begin{equation*}
G\left(\bar{x}_{0}, \bar{x}_{s}\right)=\frac{e^{-i k \hat{r}}}{4 \pi \hat{r}}, \quad \hat{r}=\left|\bar{x}_{0}-\bar{x}_{s}\right| \tag{12.1.6}
\end{equation*}
$$

The gradient operator is applied at $\bar{x}_{s}$, so that

$$
\begin{equation*}
\nabla G\left(\bar{x}_{0}, \bar{x}_{s}\right)=\frac{\bar{x}_{s}-\bar{x}_{0}}{\hat{r}} \frac{d}{d \hat{r}} G\left(\bar{x}_{0}, \bar{x}_{s}\right)=\frac{1}{4 \pi \hat{r}}\left(i k+\frac{1}{\hat{r}}\right) e^{-i k \hat{r}}\left(\frac{\bar{x}_{0}-\bar{x}_{s}}{\hat{r}}\right) \tag{12.1.7}
\end{equation*}
$$

The viewpoint in Eq. (12.1.5) is that the scattered pressure is the field radiated from $S$ by the total pressure field. An alternative form stems from the perspective of Eqs. (12.1.3) and (12.1.4), which holds that $P_{\mathrm{s}}$ is a solution of the Helmholtz equation that satisfies the Sommerfeld radiation condition. Thus, $P_{\mathrm{s}}$ by itself must be consistent with the KHIT,

$$
\begin{equation*}
P_{\mathrm{s}}\left(\bar{x}_{0}\right)=\iint_{\mathcal{S}}\left[P_{\mathrm{s}}\left(\bar{x}_{s}\right) \bar{n}\left(\bar{x}_{s}\right) \cdot \nabla G\left(\bar{x}_{0}, \bar{x}_{s}\right)-\bar{n}\left(\bar{x}_{s}\right) \cdot \nabla P_{\mathrm{s}}\left(\bar{x}_{s}\right) G\left(\bar{x}_{0}, \bar{x}_{s}\right)\right] d \mathcal{S} \tag{12.1.8}
\end{equation*}
$$

In this view, the scattered pressure at a field point is the field radiated from $\mathcal{S}$ due to the scattered pressure and particle velocity on that surface.

Some individuals are puzzled by the availability of two forms of the KHIT for $P_{\mathrm{s}}$, so let us examine this issue. By definition, the total pressure is $P=P_{\mathrm{s}}+P_{\mathrm{I}}$. Hence, if we subtract the above version from Eq. (12.1.5), we find that

$$
\begin{equation*}
\iint_{\mathcal{S}}\left[P_{\mathrm{I}}\left(\bar{x}_{s}\right) \bar{n}\left(\bar{x}_{s}\right) \cdot \nabla G\left(\bar{x}_{0}, \bar{x}_{s}\right)-\bar{n}\left(\bar{x}_{s}\right) \cdot \nabla P_{\mathrm{I}}\left(\bar{x}_{s}\right) G\left(\bar{x}_{0}, \bar{x}_{s}\right)\right] d \mathcal{S}=0 \tag{12.1.9}
\end{equation*}
$$

This identity may be derived in a different way. The incident field satisfies the KHIT for the surface $\mathcal{S}_{0}$ from which it emanates, that is,

$$
\begin{equation*}
P_{\mathrm{I}}\left(\bar{x}_{0}\right)=\iint_{\mathcal{S}_{0}}\left[P_{\mathrm{I}}\left(\bar{x}_{s}\right) \bar{n}\left(\bar{x}_{s}\right) \cdot \nabla G\left(\bar{x}_{0}, \bar{x}_{s}\right)-\bar{n}\left(\bar{x}_{s}\right) \cdot \nabla P_{\mathrm{I}}\left(\bar{x}_{s}\right) G\left(\bar{x}_{0}, \bar{x}_{s}\right)\right] d \mathcal{S} \tag{12.1.10}
\end{equation*}
$$

At the same time, we could form the KHIT for $P_{\mathrm{I}}\left(\bar{x}_{0}\right)$ by including the scattering body in the domain. Then, the surface integral would consist of contributions from $S$ and $S_{0}$. According to the preceding, the contribution from $\mathcal{S}_{0}$ gives $P_{\mathrm{I}}\left(\bar{x}_{0}\right)$, so the contribution from $\mathcal{S}$ must be zero. This is the condition stated in Eq. (12.1.9). Whether it is preferable to use Eq. (12.1.5) or Eq. (12.1.8) sometimes depends on what is known about the field.

### 12.2 Scattering by Heterogeneity

In Example 11.5 of the previous chapter, we used ray theory to examine how a signal passes through a region in which the fluid properties vary spatially. In the situation considered there, the density and sound speed were continuous functions of position, so there was no impedance mismatch. Consequently, sound was transmitted into and out of the heterogenous region without reflections. In the situation, we consider here $\mathcal{S}$ is the surface that bounds a region in which the density $\rho^{\prime}$ and sound speed $c^{\prime}$ are different from the ambient properties exterior to $\mathcal{S}$. Thus, the incident signal is scattered by $\mathcal{S}$, as well as being transmitted into the interior, then out again. Both $\rho^{\prime}$ and $c^{\prime}$ are taken to be constant, so the scattering region is inhomogeneous from the perspective of the overall domain, although it is internally homogeneous.

### 12.2.1 General Equations

We cannot determine the scattered pressure directly from the KHIT integral because we do not know the surface pressure and particle velocity. Perhaps we can gain some insight if we examine the field in the homogeneous region $\mathcal{V}$ interior to $\mathcal{S}$. A prime will denote that a quantity is associated with the interior, so the position of a point in $\mathcal{V}$ is labeled in Fig. 12.1 as $\bar{x}^{\prime}$. Point $C$ is central to $\mathcal{V}$, such as the centroid, and $\bar{x}_{0}$ is an arbitrary field point exterior to $\mathcal{V}$. The $z$-axis has any convenient reference orientation. It serves as the axial direction for a set of spherical coordinates, $r, \psi_{0}, \theta_{0}$ that locates the field point.

Fig. 12.1 Definition of positions and coordinates relative to fluid region $\mathcal{V}$ whose properties differ from those of the surrounding fluid


The pressure and normal particle velocity must be continuous along the interface $\mathcal{S}$ between the inner and outer fluids, so it must be that

$$
\begin{equation*}
P^{\prime}\left(\bar{x}_{s}\right)=P\left(\bar{x}_{s}\right) \text { and } \frac{1}{i \omega \rho^{\prime}} \bar{n}\left(\bar{x}_{s}\right) \cdot \nabla P^{\prime}\left(\bar{x}_{s}\right)=\frac{1}{i \omega \rho_{0}} \bar{n}\left(\bar{x}_{s}\right) \cdot \nabla P\left(\bar{x}_{s}\right) \tag{12.2.1}
\end{equation*}
$$

Upon substitution of these relations, Eq. (12.1.5) becomes

$$
\begin{equation*}
P_{\mathrm{S}}\left(\bar{x}_{0}\right)=\iint_{\mathcal{S}} \bar{n}\left(\bar{x}_{s}\right) \cdot\left[P^{\prime}\left(\bar{x}_{s}\right) \nabla G\left(\bar{x}_{0}, \bar{x}_{s}\right)-\frac{\rho_{0}}{\rho^{\prime}} \nabla P^{\prime}\left(\bar{x}_{s}\right) G\left(\bar{x}_{0}, \bar{x}_{s}\right)\right] d \mathcal{S} \tag{12.2.2}
\end{equation*}
$$

The divergence theorem converts the integral to one that extends over the interior domain. Note that $\bar{n}\left(\bar{x}_{s}\right)$ points out of the interior domain, so no sign change is entailed in this operation. Thus, we have

$$
\begin{equation*}
P_{\mathrm{s}}\left(\bar{x}_{0}\right)=\iiint_{\mathcal{V}} \nabla \cdot\left[P^{\prime}\left(\bar{x}^{\prime}\right) \nabla G\left(\bar{x}_{0}, \bar{x}^{\prime}\right)-\frac{\rho_{0}}{\rho^{\prime}} \nabla P^{\prime}(\bar{x}) G\left(\bar{x}_{0}, \bar{x}^{\prime}\right)\right] d \mathcal{V} \tag{12.2.3}
\end{equation*}
$$

where $\nabla$ is the gradient at interior point $\bar{x}^{\prime}$.
The fact that $\bar{x}_{0}$ is exterior to $\mathcal{V}$ simplifies this integral because nowhere does an interior point $\bar{x}^{\prime}$ coincide with $\bar{x}_{0}$. Consequently, $G\left(\bar{x}_{0}, \bar{x}^{\prime}\right)$ is a solution of the Helmholtz equation with no additional source term, that is, $\nabla^{2} G\left(\bar{x}_{0}, \bar{x}^{\prime}\right)=$ $-k^{2} G\left(\bar{x}_{0}, \bar{x}^{\prime}\right)$. In addition, $P^{\prime}\left(\bar{x}^{\prime}\right)$ is a solution of the Helmholtz equation for the interior domain, where the sound speed is $c^{\prime}$. It follows that $\nabla^{2} P^{\prime}\left(\bar{x}^{\prime}\right)=$ $-\left(\omega / c^{\prime}\right)^{2} P^{\prime}\left(\bar{x}^{\prime}\right)$. These relations convert the KHIT to

$$
\begin{align*}
P_{\mathrm{S}}\left(\bar{x}_{0}\right)= & \iiint_{\mathcal{V}}\left[\left(1-\frac{\rho_{0}}{\rho^{\prime}}\right) \nabla P^{\prime}\left(\bar{x}^{\prime}\right) \cdot \nabla G\left(\bar{x}_{0}, \bar{x}^{\prime}\right)\right. \\
& \left.-k^{2}\left(1-\frac{\rho_{0} c^{2}}{\rho^{\prime}\left(c^{\prime}\right)^{2}}\right) P^{\prime}\left(\bar{x}^{\prime}\right) G\left(\bar{x}_{0}, \bar{x}^{\prime}\right)\right] d \mathcal{V} \tag{12.2.4}
\end{align*}
$$

Further simplifications result if we limit consideration to cases where $\bar{x}_{0}$ is in the farfield. The farfield approximations apply if the smallest distance from a point within $\mathcal{V}$ to $\bar{x}_{0}$ is much greater than the largest dimension of the region. The farfield approximations use some central point $C$ as the reference point, so that $\bar{x}_{C} \equiv \overline{0}$. The length of the line from point $C$ to $\bar{x}_{0}$ is the radial distance $r$, and the direction is $\bar{e}_{r}$. The line from any point $\bar{x}^{\prime}$ to $\bar{x}_{0}$ is considered to be parallel to $\bar{e}_{r}$, so the distance between $\bar{x}^{\prime}$ and $\bar{x}_{0}$ differs from $r$ by the projection of the vector connecting those points. The relations are

$$
\begin{equation*}
r=\left|\bar{x}_{0}\right|, \quad \bar{e}_{r}=\frac{\bar{x}_{0}}{r}, \quad \hat{r} \equiv\left|\bar{x}_{0}-\bar{x}^{\prime}\right| \approx r-\bar{x}^{\prime} \cdot \bar{e}_{r} \tag{12.2.5}
\end{equation*}
$$

This leads to a simplified form of the Green's function in Eq. (12.1.6),

$$
\begin{equation*}
G\left(\bar{x}_{0}, \bar{x}^{\prime}\right) \approx \frac{e^{-i k r}}{4 \pi r} e^{i k \bar{x}^{\prime} \cdot \bar{e}_{r}} \tag{12.2.6}
\end{equation*}
$$

Furthermore, $\nabla G\left(\bar{x}_{0}, \bar{x}^{\prime}\right)$ is the gradient at $\bar{x}^{\prime}$, with $\bar{x}_{0}$ held fixed, so that

$$
\begin{equation*}
\nabla G\left(\bar{x}_{0}, \bar{x}^{\prime}\right)=\nabla \hat{r} \frac{d}{d \hat{r}} G\left(\bar{x}_{0}, \bar{x}^{\prime}\right) \approx \bar{e}_{r} \frac{i k e^{-i k r}}{4 \pi r} e^{i k \bar{x}^{\prime} \cdot \bar{e}_{r}} \tag{12.2.7}
\end{equation*}
$$

The orientation of $\bar{x}_{0}$ is described by the polar angle $\psi_{0}$ and azimuthal angle $\theta_{0}$. Hence, substitution of these expressions into Eq. (12.2.4) leads to recognition that farfield scattering from a heterogeneity is described by two directivity factors. The general representation is

$$
\begin{equation*}
P_{\mathrm{s}}\left(\bar{x}_{0}\right)=\frac{e^{-i k r}}{4 \pi r}\left[\left(1-\frac{\rho_{0}}{\rho^{\prime}}\right) F_{1}\left(\psi_{0}, \theta_{0}\right)+\left(1-\frac{\rho_{0} c^{2}}{\rho^{\prime}\left(c^{\prime}\right)^{2}}\right) F_{2}\left(\psi_{0}, \theta_{0}\right)\right] \tag{12.2.8}
\end{equation*}
$$

where the directivities are

$$
\begin{gather*}
F_{1}\left(\psi_{0}, \theta_{0}\right)=i k \bar{e}_{r} \cdot \iiint_{\mathcal{V}} e^{i k \bar{x}^{\prime} \cdot \bar{e}_{r}} \nabla P^{\prime}\left(\bar{x}^{\prime}\right) d \mathcal{V} \\
F_{2}\left(\psi_{0}, \theta_{0}\right)=-k^{2} \iiint_{\mathcal{V}} e^{i k \bar{x}^{\prime} \cdot \bar{e}_{r}} P^{\prime}\left(\bar{x}^{\prime}\right) d \mathcal{V} \tag{12.2.9}
\end{gather*}
$$

Without further development, these relations merely tell us that the strength of the scattered field depends on the mismatch of densities and of the bulk moduli, with the former being attributable to the particle velocity within the scattering volume and the latter due to the pressure in that region. Evaluation of the directivity factors requires judicious approximations of the interior field. Two that have been used consider different situations. The Born approximation is based on an assumption that the properties of the scattering region do not drastically differ from those of the surrounding homogeneous fluid. An approximation due to Rayleigh addresses objects whose size is much smaller than an acoustic wavelength.

### 12.2.2 The Born Approximation

If the region inside $\mathcal{S}$ was the same as the fluid in the outer region, the pressure inside $\mathcal{S}$ would be the incident value at that location. The Born approximation addresses the situation in which the difference between the fluids is not substantial, that is, when $\rho^{\prime} \approx \rho_{0}$ and $\rho^{\prime}\left(c^{\prime}\right)^{2} \approx \rho_{0} c^{2}$. In that case, it is reasonable to assume that the inside pressure differs little from $P_{\mathrm{I}}$. Setting $P^{\prime}\left(\bar{x}^{\prime}\right)=P_{\mathrm{I}}\left(\bar{x}^{\prime}\right)$ in Eq. (12.2.4) converts the directivity factors in Eq. (12.2.9) to computable functions.

The typical situation is a plane wave that propagates in direction $\bar{e}_{\mathrm{I}}$. Let $\hat{P}_{\mathrm{I}}$ be the amplitude at the origin, so that

$$
\begin{equation*}
P_{\mathrm{I}}\left(\bar{x}^{\prime}\right)=\hat{P}_{\mathrm{I}} \exp \left(-i k \bar{x}^{\prime} \cdot \bar{e}_{\mathrm{I}}\right), \quad \nabla P^{\prime}\left(\bar{x}^{\prime}\right)=-i k \bar{e}_{\mathrm{I}} \hat{P}_{\mathrm{I}} \exp \left(-i k \bar{x}^{\prime} \cdot \bar{e}_{\mathrm{I}}\right) \tag{12.2.10}
\end{equation*}
$$

Substitution of these representations into Eq. (12.2.9) gives

$$
\begin{gather*}
F_{1}\left(\psi_{0}, \theta_{0}, \psi_{\mathrm{I}}, \theta_{\mathrm{I}}\right)=k^{2} \hat{P}_{\mathrm{I}} \bar{e}_{r} \cdot \bar{e}_{\mathrm{I}} \mathcal{L}\left(\psi_{0}, \theta_{0}, \psi_{\mathrm{I}}, \theta_{\mathrm{I}}\right), \\
F_{2}\left(\psi_{0}, \theta_{0}, \psi_{\mathrm{I}}, \theta_{\mathrm{I}}\right)=-k^{2} \hat{P}_{\mathrm{I}} \mathcal{L}\left(\psi_{0}, \theta_{0}, \psi_{\mathrm{I}}, \theta_{\mathrm{I}}\right) \\
\mathcal{L}\left(\psi_{0}, \theta_{0}, \psi_{\mathrm{I}}, \theta_{\mathrm{I}}\right)=\iiint e^{i k\left(\bar{x}^{\prime} \cdot \bar{e}_{r}-\bar{x}^{\prime} \cdot \bar{e}_{\mathrm{I}}\right)} d \mathcal{V} \tag{12.2.11}
\end{gather*}
$$

Because the integrals for the directivity factors $F_{1}$ and $F_{2}$ extend over the domain of the scattering region, it makes sense to align the coordinate system consistently with the shape of the body. Then, $\bar{e}_{\mathrm{I}}$ and $\bar{e}_{r}$ to the field point have arbitrary orientation relative to the coordinate system. We use spherical angles $\psi_{\mathrm{I}}$ and $\theta_{\mathrm{I}}$ defined relative to the $z$-axis to describe $\bar{e}_{\mathrm{I}}$, while angles $\psi_{0}$ and $\theta_{0}$ describe $\bar{e}_{r}$.

$$
\begin{align*}
& \bar{e}_{\mathrm{I}}=\sin \psi_{\mathrm{I}} \cos \theta_{\mathrm{I}} \bar{e}_{x}+\sin \psi_{\mathrm{I}} \sin \theta_{\mathrm{I}} \bar{e}_{y}+\cos \psi_{\mathrm{I}} \bar{e}_{z}  \tag{12.2.12}\\
& \bar{e}_{r}=\sin \psi_{0} \cos \theta_{0} \bar{e}_{x}+\sin \psi_{0} \sin \theta_{0} \bar{e}_{y}+\cos \psi_{0} \bar{e}_{z}
\end{align*}
$$

The presence of $\psi_{\mathrm{I}}$ and $\theta_{\mathrm{I}}$ as variables affecting the directivity factors emphasizes that the scattered pressure depends on the propagation direction of the incident wave
relative to the body, as well as the direction to the field point, which is defined by $\psi_{0}$ and $\theta_{0}$. The position $\bar{x}^{\prime}$ and differential volume element $d \mathcal{V}$ may be described in terms of Cartesian coordinates, cylindrical coordinates, or spherical coordinates, whichever is most convenient.

This description of Born scattering is not limited to a certain frequency range. However, its utility is limited by the condition that the scattering region is homogeneous with properties that are very much like those of the surrounding fluid.

> EXAMPLE 12.1 A plane wave is incident on a region in the shape of a cylinder whose radius is $a$ and length is $L$. The sound speed and density in this region are very close to the ambient properties, so the Born approximation may be applied. Derive an expression for the scattered pressure in the form $(r / a)\left|P_{s}\left(\bar{x}_{0}\right) / \hat{P}_{\mathrm{I}}\right|$ for arbitrary $\bar{e}_{\mathrm{I}}$ and $\bar{e}_{r}$. Then, specialize this expression to the case of backscatter, for which $\bar{e}_{r}=-\bar{e}_{\mathrm{I}}$. Use that expression to evaluate the dependence of the backscattered pressure on the orientation of the incident wave for $k a=0.1,1$, and 10 . The aspect ratio is $L / a=20$, and the fluid properties are $\rho_{0} / \rho^{\prime}=c / c^{\prime}=0.95$.

## Significance

The solution demonstrates the way in which scattering integrals are formulated in general. The results that are obtained are a specific case of backscatter, which later will be addressed in greater detail. Most importantly, the results are illustrative of general scattering properties.

## Solution

The starting point is Fig. 1, which illustrates the scattering region, the unit vectors, and the $x y z$ coordinate system for the description of vectors. The origin is situated at the centroid of the cylinder, and the $z$-axis is aligned with the centerline. The axisymmetry of the system makes it permissible to situate the $x$-axis in the plane containing $\bar{e}_{\mathrm{I}}$ and $\bar{e}_{z}$. This means that we can set $\theta_{\mathrm{I}}=0$ without losing generality of the analysis.


Figure 1

The shape is best described in terms of cylindrical coordinates defined relative to $x y z$, so that $d \mathcal{V}=R d \theta d R d z$. The relevant vectors are

$$
\begin{gather*}
\bar{e}_{\mathrm{I}}=\sin \psi_{\mathrm{I}} \bar{e}_{x}+\cos \psi_{\mathrm{I}} \bar{e}_{z} \\
\bar{e}_{r}=\sin \psi_{0} \cos \theta_{0} \bar{e}_{x}+\sin \psi_{0} \sin \theta_{0} \bar{e}_{y}+\cos \psi_{0} \bar{e}_{z}  \tag{1}\\
\bar{x}^{\prime}=R \cos \theta \bar{e}_{x}+R \sin \theta \bar{e}_{y}+z \bar{e}_{z}
\end{gather*}
$$

When these representations are substituted into the integral in Eq. (12.2.11), the result is

$$
\begin{aligned}
\mathcal{L}\left(\psi_{0}, \theta_{0}, \psi_{\mathrm{I}}, \theta_{\mathrm{I}}\right)= & \int_{-L / 2}^{L / 2} \int_{0}^{a} \int_{-\pi}^{\pi} \exp \left[i k R\left(\sin \psi_{0} \cos \theta_{0}-\sin \psi_{\mathrm{I}}\right) \cos \theta\right. \\
& \left.+i k R\left(\sin \psi_{0} \sin \theta_{0}\right) \sin \theta+i k z\left(\cos \psi_{0}-\cos \psi_{\mathrm{I}}\right)\right] R d \theta d R d z
\end{aligned}
$$

Several simplifications are available. The first two terms in the integrand are independent of $z$, whereas the last term is independent of $R$ and $\theta$. The integral may be factorized accordingly. Furthermore, the integral over $z$ may be evaluated analytically. The integral over the range of $\theta$ may be split into integrals over positive and negative $\theta$. Replacing $\theta$ by $-\tilde{\theta}$ in the latter interval allows the integral to extend only over positive $\theta$. The result is that

$$
\begin{equation*}
\mathcal{L}\left(\psi_{0}, \theta_{0}, \psi_{\mathrm{I}}, \theta_{\mathrm{I}}\right)=a^{2} L \mathcal{G}\left(\psi_{0}, \theta_{0}, \psi_{\mathrm{I}}\right) \mathcal{H}\left(\psi_{0}, \psi_{\mathrm{I}}\right) \tag{2}
\end{equation*}
$$

The $\mathcal{H}$ factor is the result of integration over $z$,

$$
\mathcal{H}\left(\psi_{0}, \psi_{\mathrm{I}}\right)=\int_{-1 / 2}^{1 / 2} \exp \left[i k L \hat{z}\left(\cos \psi_{0}-\cos \psi_{\mathrm{I}}\right)\right] d \hat{z}
$$

It may be integrated analytically, which leads to

$$
\mathcal{H}\left(\psi_{0}, \psi_{\mathrm{I}}\right)=\left\{\begin{array}{l}
\frac{\sin \left[(k L / 2)\left(\cos \psi_{0}-\cos \psi_{\mathrm{I}}\right)\right]}{(k L / 2)\left(\cos \psi_{0}-\cos \psi_{\mathrm{I}}\right)} \text { if } \psi_{0} \neq \psi_{\mathrm{I}}  \tag{3}\\
1 \text { if } \psi_{0}=\psi_{\mathrm{I}}
\end{array}\right.
$$

The $\mathcal{G}$ factor comes from integration over $\theta$ and $R$. It is

$$
\begin{align*}
\mathcal{G}\left(\psi_{0}, \theta_{0}, \psi_{\mathrm{I}}\right)= & 2 \int_{0}^{1} \int_{0}^{\pi} \exp \left[i k a \hat{R}\left(\sin \psi_{0} \cos \theta_{0}-\sin \psi_{\mathrm{I}}\right) \cos \theta\right]  \tag{4}\\
& \times \cos \left[k a \hat{R}\left(\sin \psi_{0} \sin \theta_{0}\right) \sin \theta\right] \hat{R} d \theta d \hat{R}
\end{align*}
$$

Numerical methods are required to evaluate this function for arbitrary orientation angles. However, a closed form result is available for the backscatter case.

After the value of $\mathcal{L}\left(\psi_{0}, \theta_{0}, \psi_{\mathrm{I}}, \theta_{\mathrm{I}}\right)$ is determined for a specific set of angles, the directivity factors in Eq. (12.2.11) are known. Their substitution, and $\bar{e}_{r} \cdot \bar{e}_{\mathrm{I}}$ from Eq. (1), into Eq. (12.2.8) gives the scattered pressure. Collecting the parameters into nondimensional groups leads to

$$
\begin{align*}
P_{\mathrm{S}}\left(\bar{x}_{0}\right)= & \frac{\hat{P}_{\mathrm{I}}}{4 \pi}\left(\frac{a}{r}\right) e^{-i k r}(k a)(k L)\left[\left(1-\frac{\rho_{0}}{\rho^{\prime}}\right)\left(\sin \psi_{0} \cos \theta_{0} \sin \psi_{\mathrm{I}}+\cos \psi_{0} \cos \psi_{\mathrm{I}}\right)\right. \\
& \left.-\left(1-\frac{\rho_{0} c^{2}}{\rho^{\prime}\left(c^{\prime}\right)^{2}}\right)\right] \mathcal{G}\left(\psi_{0}, \theta_{0}, \psi_{\mathrm{I}}\right) \mathcal{H}\left(\psi_{0}, \psi_{\mathrm{I}}\right) \tag{5}
\end{align*}
$$

Evaluation of the $\mathcal{G}$ function for arbitrary $\psi_{\mathrm{I}}, \psi_{0}$, and $\theta_{0}$ requires numerical integration, which is easier if the available software has a routine for double integration. In MATLAB, it is dblquad, whose input arguments are the name of a function that evaluates the integrand, as well as the lower and upper integration limits for both variables.

Equation (5) offers little insight to the nature of the scattered field. The results for an arbitrary set of angles could be displayed as a spherical plot for fixed $\psi_{\mathrm{I}}$, in which $\psi_{0}$ and $\theta_{0}$ are the spherical angles and $(r / a)\left|P_{s} / \hat{P}_{\mathrm{I}}\right|$ is the radial distance. Many such plots would describe the dependence on $\psi_{\mathrm{I}}$. Fortunately, the information that is requested sets $\bar{e}_{r}=-\bar{e}_{\mathrm{I}}$. This corresponds to setting $\theta_{0}=\pi$ and $\psi_{0}=\pi-\psi_{\mathrm{I}}$. For these angles, the preceding expression reduces to

$$
\begin{gather*}
\left(\frac{r}{a}\right) \frac{P_{\mathrm{s}}\left(\bar{x}_{0}\right)}{\hat{P}_{\mathrm{I}}}=\frac{e^{-i k r}}{4 \pi}(k a)(k L)\left[\frac{\rho_{0}}{\rho^{\prime}}\left(1+\frac{c^{2}}{\left(c^{\prime}\right)^{2}}\right)-2\right]  \tag{6}\\
\mathcal{G}\left(\pi-\psi_{\mathrm{I}}, \pi, \psi_{\mathrm{I}}\right) \mathcal{H}\left(\pi-\psi_{\mathrm{I}}, \psi_{\mathrm{I}}\right)
\end{gather*}
$$

where

$$
\begin{gathered}
\mathcal{G}\left(\pi-\psi_{\mathrm{I}}, \pi, \psi_{\mathrm{I}}\right)=2 \int_{0}^{1} \int_{0}^{\pi} \exp \left(-2 i k a \hat{R} \sin \psi_{\mathrm{I}} \cos \theta\right) \hat{R} d \theta d \hat{R} \\
\mathcal{H}\left(\pi-\psi_{\mathrm{I}}, \psi_{\mathrm{I}}\right)=\frac{\sin \left(k L \cos \psi_{\mathrm{I}}\right)}{k L \cos \psi_{\mathrm{I}}}
\end{gathered}
$$

Analytical evaluation is possible by invoking general formulas provided in Abramowitz and Stegun's compendium. ${ }^{1}$ The eventual result is

$$
\begin{equation*}
\mathcal{G}\left(\pi-\psi_{\mathrm{I}},-\pi, \psi_{\mathrm{I}}\right)=\pi \frac{J_{1}\left(2 k a \sin \psi_{\mathrm{I}}\right)}{k a \sin \psi_{\mathrm{I}}} \tag{7}
\end{equation*}
$$

There are different ways in which parametric trends may be viewed. We can consider the value of $P_{s}$ that would be measured at a fixed location, which sets $r$.

[^56]Because $a^{2} L=\mathcal{V} / \pi$, that view considers $P_{s}$ to be proportional to the volume of the heterogenous medium, as well as the square of the frequency. However, changing the dimensions or the frequency also has a significant effect on the angular dependence. The first of Eq. (6) indicates that $\mathcal{G}$ is a function of $k a$, as well as $\psi_{\mathrm{I}}$. The value of $\mathcal{G}$ for $\psi_{\mathrm{I}}=0$ is $\pi$ for any $k a$, as well as for any $\psi_{\mathrm{I}}$ if $k a \ll 1$. Furthermore, $\mathcal{G}=\pi$ is the maximum value for any $k a$ and $\psi_{\mathrm{I}}$. These properties mean that the $\mathcal{G}$ function leads to enhanced directivity in the axial direction as $k a$ increases. In the case of $\mathcal{H}$, its maximum is $\mathcal{H}=1$. This value occurs at any $\psi_{\mathrm{I}}$ if $k L \ll 1$, and at $\psi_{\mathrm{I}}=\pi / 2$ for any $k L$. Thus, $\mathcal{H}$ leads to greater directivity in the transverse direction as the frequency increases. The low-frequency limit is examined in detail in the next section.

Figure 2 displays polar plots of $(r / a)\left|P_{s} / P_{\mathrm{I}}\right|$ at the designated frequencies. Backscatter at broadside incidence is somewhat larger at the lowest frequency, and it is strongly directional in the broadside direction at the middle frequency. At the highest frequency, backscatter is comparable for end-on and broadside incidence, and both are confined to a smaller range of angles.


Figure 2

To gain insight to this shifting trend, let us consider individually $\mathcal{G}\left(\pi-\psi_{\mathrm{I}}\right.$, $\left.-\pi, \psi_{\mathrm{I}}\right)$ and $\mathcal{H}\left(\pi-\psi_{\mathrm{I}}, \psi_{\mathrm{I}}\right)$, whose product describes the backscatter directivity. Examination of Eq. (7) reveals that the maximum value of the former is $\mathcal{G}=\pi$ at end-on incidence and that $\mathcal{G}=\pi J_{1}(2 k a) /(k a)$ at broadside incidence. This value decreases in an oscillatory manner with increasing $k a$. In contrast, the maximum value of $\mathcal{H}$ broadside is $\mathcal{H}=1$, and $\mathcal{H}=\sin (k L) /(k L)$ for end-on incidence. This too decreases in an oscillatory manner with increasing frequency. With these trends in mind, consider the graphs in Fig. 3. Each function is large in a narrow interval. The directivity of $P_{s}$ is the product of the functions, which explains why it is very small away from $\psi_{\mathrm{I}}=0^{\circ}, 90^{\circ}$, and $180^{\circ}$. Increasing $k a$ increases $k L$, so $J_{1}(2 k a)$ and $\sin (k L)$ will grow and decrease. If either term is near zero, the value of $p_{s}$ will be very small in the respective direction. Consequently, whether backscatter is greater at broadside or end-on incidence strongly depends on the frequency.


Figure 3

### 12.3 Rayleigh Scattering Limit

The Rayleigh scattering limit is a general term that refers to the field scattered by any object whose size is much less than the acoustic wavelength. Specifically, if $a$ is the largest dimension, then the Rayleigh scattering limit is the behavior when $k a \ll 1$. This condition leads to several analytical simplifications. It also will disclose some qualitative aspects of the field that are not evident from the general equations.

The Rayleigh limit is sometimes referred to as a quasistatic approximation. This terminology stems from the fact that because of the small size of the body, the phase of any long-wavelength signal will be essentially constant over the extent of the body. This means that any complex exponential representing a phase shift within the scattering body or on its surface may be approximated as having unit value.

### 12.3.1 The Rayleigh Limit of the Born Approximation

The Born approximation is not limited by the size of the heterogeneous region. Its associated directivity factors must be evaluated for a specific body, and a specific frequency. The Rayleigh limit of the Born approximation yields a formula for the scattered field that does not require evaluation of integrals. At the same time, it is less general because it considers the scattering region to be much smaller than a wavelength.

The previous example indicates how we should proceed to obtain the low frequency version. As before, we shall only consider the case of an incident plane wave, which is described by Eq. (12.2.10). Let $a$ be the largest dimension of the scattering region. If the origin is interior to the region, then any position within this region is such that $\left|\bar{x}^{\prime}\right| \leq a$. Furthermore, we are interested in the situation when $2 \pi / k \gg a$. It follows that $k\left|\bar{x}^{\prime}\right| \ll 1$, so that the integrand in Eq. (12.2.11) is essentially one, which gives $\mathcal{L}=\mathcal{V}$. The corresponding directivity factors are
$F_{1}=k^{2} \hat{P}_{\mathrm{I}} \bar{e}_{r} \cdot \bar{e}_{\mathrm{I}} \mathcal{V}$ and $F_{2}=-k^{2} \hat{P}_{\mathrm{I}} \mathcal{V}$. The only directional factor that arises is $\bar{e}_{r} \cdot \bar{e}_{\mathrm{I}}$. Therefore, we may define the $z$-axis to be parallel to $\bar{e}_{\mathrm{I}}$. This reduces Eq. (12.2.8) to

$$
\begin{equation*}
P_{s}\left(\bar{x}_{0}\right)=k^{2} \hat{P}_{\mathrm{I}} \frac{e^{-i k r}}{4 \pi r} \mathcal{V}\left[\left(\frac{\rho_{0} c^{2}}{\rho^{\prime}\left(c^{\prime}\right)^{2}}-1\right)-\left(\frac{\rho_{0}}{\rho^{\prime}}-1\right) \cos \psi_{0}\right] \tag{12.3.1}
\end{equation*}
$$

where $\psi_{0}$ is the angle between $\bar{e}_{\mathrm{I}}$ and the radial direction to the field point. This expression tells us that the strength of the scattered field at a fixed location is proportional to the volume of the scatterer. This attribute surprises some individuals who intuitively think that the cross-sectional area represents the amount the sound is blocked. Another important observation is that the field consists of a monopole whose strength is proportional to the difference of the bulk moduli, combined with a dipole component that is proportional to the difference of densities. The polar angle to the field point is measured from the incident direction $\bar{e}_{\mathrm{I}}$. Thus, the dipole is aligned in the direction of the incident wave. Both the monopole and dipole strengths are proportional to the square of the frequency.

It is instructive to compare this result to the solution of Example 12.1, which is specific to a cylindrical region. That result is valid for any frequency, but it is only descriptive of a specific shape. The Rayleigh limit described Eq. (12.3.1) is valid only for low frequencies, but it is descriptive of scattering from any region, regardless of its shape. Another difference is that the example was limited to situations where the scattered signal is observed on the same line as that from the planar source to the region. This corresponds to $\psi_{\mathrm{I}}=180^{\circ}$ in the above expression. The result for a cylindrical region when $k a \ll 1$ is identical to Eq. (12.3.1) when $\mathcal{V}=\pi a^{2} L$ and $\cos \psi_{\mathrm{I}}=-1$. Despite the generality of the present result, the assumptions embedded in the model lead to a description of the scattered pressure that depends on the location of the field point, but is independent of the orientation of the body. In contrast, such dependence is displayed in the Born approximation.

Viewing the dependence of the scattered pressure in Eq. (12.3.1) in terms of nondimensional variables offers a different perspective. The largest dimension has been denoted as $a$, so the volume can be described as $\beta a^{3}$, where $\beta$ is a shape factor. Then $r / a\left|P_{s} / \hat{P}_{\mathrm{I}}\right|$ depends solely on $\beta,(k a)^{2}$, and $\psi_{\mathrm{I}}$.

This trend was used by Rayleigh ${ }^{2}$ to explain why the sky is blue. At midday, the light rays that are incident at our location from the sun essentially are white, which means that they are composed of all colors in the visual spectrum in nearly equal amounts. These rays constitute the image of the sun. If there were no scattering, these are the only rays that would travel from the sun to an observer on the ground. The rest of the sky would appear to be black, as it does on the Moon. What actually happens is that the light rays that propagate through the atmosphere in other directions are scattered by dust and water particles in the atmosphere. Some of these rays reach an observer on the ground. The Rayleigh description of scattering indicates that the

[^57]blue spectrum, which comprises the short-wavelength/high-frequency portion of the visual spectrum, scatters more strongly than the red spectrum. Thus, the contribution of the blue spectrum to what people see when they view the sky is stronger than the contribution of the red spectrum.

### 12.3.2 Mismatched Heterogeneous Region

If the properties of a small region of fluid differ significantly from those of the surrounding fluid, the Born approximation is not valid. A different analysis, also based on $k a$ being small, is available in that case. Let us multiply the Helmholtz equation by $a^{2}$ and define a nondimensional gradient $\hat{\nabla}=a \nabla$. Smallness of $k a$ means that $\hat{\nabla} P$ is the dominant term in the Helmholtz equation within the heterogeneity. Hence, the Helmholtz equation for the interior field reduces to the Laplace equation. Smallness approximations do not apply to the exterior scattered field, which must satisfy the Helmholtz equation. The boundary conditions on the surface $\mathcal{S}$, which is the interface between the media, are continuity of pressure and normal particle velocity. Let $P^{\prime}(x)$ be the pressure within the heterogeneity, and $P_{\mathrm{S}}(\bar{x})$ be the scattered field. The governing equations are

$$
\left.\begin{array}{c}
\nabla^{2} P^{\prime}=0, \quad \bar{x} \in \mathcal{V} \\
\nabla^{2} P_{\mathrm{s}}+k^{2} P_{\mathrm{s}}=0, \quad \bar{x} \notin \mathcal{V} \\
P^{\prime}=P_{\mathrm{s}}+P_{\mathrm{I}}  \tag{12.3.2}\\
\frac{\rho_{0}}{\rho^{\prime}} \bar{n} \cdot \nabla P^{\prime}=\bar{n} \cdot\left(\nabla P_{\mathrm{s}}+\nabla P_{\mathrm{I}}\right)
\end{array}\right\} \bar{x} \in \mathcal{S}
$$

In addition the scattered field must satisfy the radiation condition in Eq. (12.1.4).
To see how the limitation to small $k a$ simplifies the analysis, we shall consider a specific situation. Suppose $\mathcal{V}$ is a sphere and the incident signal is a plane wave propagating in the axial direction, so that $\bar{e}_{\mathrm{I}}=\bar{e}_{z}$. Let $\hat{r}=r / a$ be the nondimensional radial position, and set $\bar{n}=\bar{e}_{r}$. Then, the incident pressure and its normal derivative on the surface of the sphere are

$$
\begin{gather*}
\left.\frac{\partial P_{\mathrm{I}}}{\partial \hat{r}}\right|_{\hat{r}=1} ^{\left.P_{\mathrm{I}}\right|_{\hat{r}=1}=\hat{P}_{\mathrm{I}} e^{-i k a \cos \psi}}=-i k a \cos \psi \hat{P}_{\mathrm{I}} e^{-i k a \cos \psi} \tag{12.3.3}
\end{gather*}
$$

Smallness of $k a$ allows us to expand the complex exponential in a series that is truncated at the lowest order frequency-dependent term, so that

$$
\begin{gather*}
\left.P_{\mathrm{I}}\right|_{\hat{r}=1}=(1-i k a \cos \psi) \hat{P}_{\mathrm{I}} \\
\left.\frac{\partial P_{\mathrm{I}}}{\partial \hat{r}}\right|_{\hat{r}=1}=-i k a \cos \psi \hat{P}_{\mathrm{I}} \tag{12.3.4}
\end{gather*}
$$

We may use a spherical harmonic expansions to represent $P_{\mathrm{s}}$ and $\tilde{P}$. Both series may be truncated at the $m=1$ harmonic, because that is the highest harmonic appearing in the incident pressure and pressure gradient on the surface, both of which constitute the excitation. The series for the scattered pressure is

$$
\begin{equation*}
P_{s}=\sum_{m=0}^{1} \hat{B}_{m} P_{m}(\cos \theta) h_{m}^{(2)}(k a \hat{r})=\hat{B}_{0} h_{0}^{(2)}(k a \hat{r})+\hat{B}_{1} h_{1}^{(2)}(k a \hat{r}) \cos \psi \tag{12.3.5}
\end{equation*}
$$

This description is valid everywhere, but it may be simplified in the vicinity of the sphere based on $k a \ll 1$ and $\hat{r}=O(1)$. This form is

$$
\begin{align*}
P_{s} & \approx \frac{B_{0}}{\hat{r}^{2}}+\frac{B_{1}}{\hat{r}^{2}} \cos \psi, \quad k a \hat{r} \ll 1 \\
B_{0} & =\frac{i}{k a} \hat{B}_{0}, \quad B_{1}=\frac{i}{2(k a)^{2}} \hat{B}_{0} \tag{12.3.6}
\end{align*}
$$

The value of $k a \hat{r}$ is much less than one everywhere within $\mathcal{V}$, so the spherical harmonic representation of the pressure in the scattering region may be represented by a similar form,

$$
\begin{equation*}
P^{\prime}=\sum_{m=0}^{\infty} A_{m} P_{m}(\cos \theta) j_{m}(k r) \approx A_{0}+A_{1} \hat{r} \cos \psi \tag{12.3.7}
\end{equation*}
$$

These series are substituted into the boundary conditions in Eq. (12.3.2), and like dependencies on $\psi$ are matched. This operation yields

$$
\begin{gather*}
A_{0}=\hat{P}_{\mathrm{I}}+B_{0} \\
A_{1}=-i k a \hat{P}_{\mathrm{I}}+B_{1} \\
0=-B_{0}  \tag{12.3.8}\\
\frac{\rho_{0}}{\rho^{\prime}} A_{1}=-i k a \hat{P}_{\mathrm{I}}-2 B_{1}
\end{gather*}
$$

Solution of these equations leads to an explicit description of the pressure and its gradient within $\mathcal{V}$,

$$
\begin{gather*}
P^{\prime}=\hat{P}_{\mathrm{I}}\left(1-\frac{3 i k}{2+\rho_{0} / \rho^{\prime}} r \cos \psi\right)  \tag{12.3.9}\\
\nabla P^{\prime}=-\frac{3 i k}{2+\rho_{0} / \rho^{\prime}} \hat{P}_{\mathrm{I}} r \cos \psi
\end{gather*}
$$

These expressions are used to evaluate the volume integrals in Eq. (12.2.9). The result is

$$
\begin{equation*}
P_{\mathrm{S}}\left(\bar{x}_{0}\right)=\hat{P}_{\mathrm{I}} k^{2} \frac{e^{-i k r}}{4 \pi r} \frac{4 \pi a^{3}}{3}\left[\left(\frac{\rho_{0} c^{2}}{\rho^{\prime}\left(c^{\prime}\right)^{2}}-1\right)-\frac{3\left(\rho_{0} / \rho^{\prime}-1\right)}{2+\rho_{0} / \rho^{\prime}} \cos \psi_{0}\right] \tag{12.3.10}
\end{equation*}
$$

The above has several features that are like the behavior predicted by the Born approximation. The magnitude of the scattered pressure increases as the square of the frequency. It consists of a monopole whose strength is proportional to the difference of the bulk moduli, and a dipole whose strength is proportional to the difference of densities. The amplitude of the monopole term matches the result of the Born approximation, Eq. (12.3.1), but the dipole moment is different. In the case where $\rho_{0} / \rho^{\prime} \approx 1$, the dipole moment derived here is $50 \%$ greater than the previous result. The Born approximation was introduced empirically, whereas the present result stems from an analysis that only assumes that $k a$ is sufficiently small.

Although the above formula was derived for a sphere, we know that at very low frequencies, geometric details that are much smaller than the wavelength are unimportant. Thus, it is reasonable to use Eq. (12.3.10) with $(4 / 3) \pi a^{3}$ replaced by $\mathcal{V}$, rather than Eq. (12.3.1), to describe scattering from a small embedded region of fluid, regardless of its shape.

### 12.3.3 Scattering from a Rigid Body

Analysis of the field scattered by an immovable body is quite complicated if the wavelength is comparable to the size of the body. Restricting our consideration to situations where the body is small enables us characterize the general nature of the scattered field. The development is founded on the Kirchhoff-Helmholtz integral theorem in Eq. (12.1.8), which states that $P_{\mathrm{s}}$ is the field radiated from $\mathcal{S}$ due to the scattered pressure and particle velocity on that surface. We begin by considering the situation where the rigid body is immobilized by some means. After that, we will formulate a description that accounts for the effect of removing restraints against movement.

## Stationary Bodies

When the surface $\mathcal{S}$ is stationary, the normal component of the particle velocity must vanish everywhere on $\mathcal{S}$. The total signal is the sum of the incident and scattered parts, so it must be that

$$
\begin{equation*}
\bar{n} \cdot \bar{V}_{\mathrm{s}}=-\bar{n} \cdot \bar{V}_{\mathrm{I}}, \quad \bar{x} \in \mathcal{S} \tag{12.3.11}
\end{equation*}
$$

Specific results may be obtained for low-frequency scattering of an incident plane wave,

$$
\begin{equation*}
P_{\mathrm{I}}=\hat{P}_{\mathrm{I}} e^{-i k \bar{e}_{\mathrm{I}} \cdot \bar{x}} \Longrightarrow \bar{V}_{\mathrm{I}}=\bar{e}_{\mathrm{I}} \frac{\hat{P}_{\mathrm{I}}}{\rho_{0} c} e^{-i k \bar{e}_{\mathrm{I}} \cdot \bar{x}} \tag{12.3.12}
\end{equation*}
$$

where $\bar{e}_{\mathrm{I}}$ is the propagation direction. It will be useful for the analysis if we situate the origin at the geometric centroid $C$ of the body, and set the phase lag of the wavefront through point $C$ to zero. This is the arrangement depicted in Fig. 12.2.

Fig. 12.2 Plane wave is incident on a rigid body


The definition of the low-frequency range is that the wavelength is much larger than the size of the scattering body. In that case, it must be true that $k|\bar{x}| \ll 1$ for any point in the body. Consequently, $\bar{V}_{\mathrm{I}}$ will be adequately described by the first few terms in a Taylor series,

$$
\begin{equation*}
\bar{V}_{\mathrm{I}}=\bar{e}_{\mathrm{I}} \frac{\hat{P}_{\mathrm{I}}}{\rho_{0} c}\left[1-i k \bar{e}_{\mathrm{I}} \cdot \bar{x}+\cdots\right] \tag{12.3.13}
\end{equation*}
$$

The surface normal velocity in the scattered signal must cancel the incident contribution, so we have

$$
\begin{equation*}
\bar{n} \cdot \bar{V}_{\mathrm{s}}=\bar{n} \cdot \bar{e}_{\mathrm{I}} \frac{\hat{P}_{\mathrm{I}}}{\rho_{0} c}\left[-1+i k \bar{e}_{\mathrm{I}} \cdot \bar{x}+\cdots\right] ; \quad \bar{x} \in \mathcal{S} \tag{12.3.14}
\end{equation*}
$$

A useful way to interpret this expression is that the scattered pressure is equivalent to the pressure radiated by the body when the normal velocity on the surface is the quantity on the right side. This viewpoint allows us to invoke the multipole expansion of radiation from a rigid body, which was the topic in Sect. 7.4.2. The first contribution to $\bar{n} \cdot \bar{V}_{\mathrm{s}}$ is the component of the particle velocity $\bar{V}_{\mathrm{I}}$ in the direction of the inward normal. In a rigid body translation, all points have the same velocity. Therefore, the first term represents a rigid body translational oscillation whose complex velocity amplitude is $-\bar{V}_{\mathrm{I}}$. Such a motion leads to dipole radiation, but does not contribute to the monopole amplitude. The second term in Eq. (12.3.14) may be written as $i \bar{n} \cdot \bar{V}_{\mathrm{I}}\left(\bar{e}_{\mathrm{I}} \cdot k \bar{x}\right)$. This term contributes to both the monopole amplitude and dipole moment. However, $|\bar{x}|$ can be no larger than the largest dimension $a$, and we have restricted our consideration to situations where $k a \ll 1$. This observation makes it permissible to neglect the contribution of this term to the dipole moment. Hence, by these considerations, we have recognized that the monopole amplitude $A$ accounts for the volume velocity associated with a normal velocity that is $i \bar{n} \cdot \bar{V}_{\mathrm{I}}\left(\bar{e}_{\mathrm{I}} \cdot k \bar{x}\right)$, whereas $\bar{D}$ is the dipole moment for a rigid body oscillation whose complex amplitude is $-\bar{V}_{\mathrm{I}}$.

The multipole expansion is given in Eq. (7.4.23). We let $G\left(\bar{x}_{0}, \bar{x}_{C}\right)$ be the freespace Green's function and ignore higher order terms on the basis that they are weak if $2 \pi / k$ is much larger than the body's size. The corresponding expression for the scattered pressure in the farfield is

$$
\begin{equation*}
P\left(\bar{x}_{0}\right)=\left(A+i k \bar{D} \cdot \bar{e}_{r}\right) \frac{e^{-i k r}}{r} \tag{12.3.15}
\end{equation*}
$$

The monopole amplitude is stated by the first of Eq. (7.4.24) to be

$$
\begin{align*}
A= & \frac{i \omega \rho_{0}}{4 \pi} \iint_{\mathcal{S}}\left[\left(\bar{n} \cdot \bar{e}_{\mathrm{I}} \frac{\hat{P}_{\mathrm{I}}}{\rho_{0} c}\right)\left(i k \bar{e}_{\mathrm{I}} \cdot \bar{\xi}\right)\right] d \mathcal{S}=-\frac{k^{2} \hat{P}_{\mathrm{I}}}{4 \pi} \iint_{\mathcal{S}}\left(\bar{n} \cdot \bar{e}_{\mathrm{I}}\right)\left(\bar{e}_{\mathrm{I}} \cdot \bar{\xi}\right) d \mathcal{S} \\
& =-\frac{k^{2} \hat{P}_{\mathrm{I}}}{4 \pi} \iiint_{\mathcal{V}} \nabla \cdot\left[\bar{e}_{\mathrm{I}}\left(\bar{e}_{\mathrm{I}} \cdot \bar{\xi}\right)\right] d \mathcal{V} \tag{12.3.16}
\end{align*}
$$

where the divergence theorem is applied with no sign change because $\bar{n}(\bar{\xi})$ is oriented outward from $\mathcal{V}$. To evaluate the gradient, let $\ell_{j}$ denote the direction cosines of $\bar{e}_{\mathrm{I}}$, so $\bar{e}_{\mathrm{I}} \cdot \bar{\xi}=\ell_{1} x+\ell_{2} y+\ell_{3} z$. Because $\bar{e}_{\mathrm{I}}$ is a constant, the integrand reduces to $\bar{e}_{\mathrm{I}} \nabla\left(\bar{e}_{\mathrm{I}} \cdot \bar{\xi}\right) \equiv \bar{e}_{\mathrm{I}} \cdot \bar{e}_{\mathrm{I}}=1$. Correspondingly, the monopole amplitude reduces to

$$
\begin{equation*}
A=-\frac{k^{2} \mathcal{V}}{4 \pi} \hat{P}_{\mathrm{I}} \tag{12.3.17}
\end{equation*}
$$

Determination of the dipole moment does not require evaluation of a surface integral. The requisite analysis has been performed in Sect. 7.4.2, which treated the field radiated by an oscillating rigid body. Thus, we may employ Eq. (7.4.39), with the translational velocity sets at $-\bar{V}_{\mathrm{I}} \equiv-\bar{e}_{\mathrm{I}} P_{1} /\left(\rho_{0} c\right)$. This equation is written in matrix form to facilitate describing the possibility that the dipole moment might not be aligned with the translational velocity. Thus, we find that

$$
\begin{equation*}
\{D\}=\frac{i k}{4 \pi} P_{\mathrm{I}}[\mathcal{V}[I]+[W]]\left\{e_{\mathrm{I}}\right\} \tag{12.3.18}
\end{equation*}
$$

The quantity $\rho_{0}[W]$ is the virtual mass, defined in Eq. (7.4.35).
The scattered pressure is obtained by adding the monopole and dipole contributions according to Eq. (12.3.15). In keeping with the matrix form of the relation for the dipole moment, this pressure field is

$$
\begin{equation*}
P_{s}\left(\bar{x}_{0}\right)=-\frac{k^{2} \hat{P}_{\mathrm{I}}}{4 \pi}\left[\mathcal{V}-\left\{e_{r}\right\}^{\mathrm{T}}[\mathcal{V}[I]+[W]]\left\{e_{\mathrm{I}}\right\}\left(1-\frac{i}{k r}\right)\right] \frac{e^{-i k r}}{r} \tag{12.3.19}
\end{equation*}
$$

where $r$ is the distance from the centroid of the scattering body to the field point. The farfield version is obtained by dropping the $i / k r$ term. The corresponding description of the force exerted by the scattered pressure is stated in Eq. (7.4.35). Replacing $\bar{V}_{C}$ in that expression with $\bar{e}_{\mathrm{I}} P_{\mathrm{I}} /\left(\rho_{0} c\right)$ gives

$$
\begin{equation*}
\{F\}=i k \hat{P}_{\mathrm{I}}[W]\left\{e_{\mathrm{I}}\right\} \tag{12.3.20}
\end{equation*}
$$

Both $\mathcal{V}$ and $[W]$ are shape-dependent quantities that scale as the cube of the size of the body, which is measured by dimension $a$. Therefore, the scattered pressure observed at a fixed field point depends on the cube of the size, the shape of the body, the square of the frequency, and the orientation of $\bar{e}_{\mathrm{I}}$ and $\bar{e}_{r}$ relative to the body. Nondimensionally, it is proportional to $\left|\hat{P}_{\mathrm{I}}\right|, a / r$, and $(k a)^{2}$, and the direction angles for both directions. These attributes, as well as the fact that the field is a superposition of a monopole and dipole, are the same as those found for Rayleigh scattering from a region of heterogeneity.

Evaluation of the scattered pressure when $[\Lambda]$ is known requires a description of $\left\{e_{\mathrm{i}}\right\}$ and $\left\{e_{r}\right\}$. For this, we may use spherical angles defined relative to the $z$-axis, with $\psi_{\mathrm{I}}$ and $\psi_{r}$ being the respective polar angles. The representation is

$$
\begin{gather*}
\left\{e_{\mathrm{I}}\right\}=\left[\begin{array}{lll}
\sin \psi_{\mathrm{I}} \cos \theta_{\mathrm{I}} & \sin \psi_{\mathrm{I}} \sin \theta_{\mathrm{I}} & \cos \psi_{\mathrm{I}}
\end{array}\right]^{\mathrm{T}}  \tag{12.3.21}\\
\left\{e_{r}\right\}=\left[\begin{array}{lll}
\sin \psi_{r} \cos \theta_{r} & \sin \psi_{r} \sin \theta_{r} & \cos \psi_{r}
\end{array}\right]^{\mathrm{T}}
\end{gather*}
$$

The expression for the scattered pressure simplifies considerably when the body is axisymmetric. We shall designate $z$ as this axis. A consequence of axisymmetry is that we may take $\bar{e}_{\mathrm{I}}$ to lie in the $z x$, so that $\theta_{\mathrm{I}}=0$. Setting the off-diagonal terms of $[\Lambda$ ] to zero reduces the farfield scattered pressure in Eq. (12.3.19) to

$$
\begin{align*}
P_{\mathrm{S}}\left(\bar{x}_{0}\right)= & -\frac{k^{2} \hat{P}_{\mathrm{I}}}{4 \pi}\left[\mathcal{V}-\left(\mathcal{V}+W_{1,1}\right) \sin \psi_{\mathrm{I}} \sin \psi_{r} \cos \theta_{r}\right. \\
& \left.-\left(\mathcal{V}+W_{3,3}\right) \cos \psi_{\mathrm{I}} \cos \psi_{r}\right] \frac{e^{-i k r}}{r} \tag{12.3.22}
\end{align*}
$$

For a sphere, $W_{1,1}=W_{3,3}=(2 / 3) \pi a^{3}$, which corresponds to the added mass being half the mass of the displaced fluid. For a thin disk, $W_{1,1}=0$ and $W_{3,3}=(8 / 3) a^{3}$. In the thin disk approximation, the displaced volume is essentially zero, so the added mass is not related to the displaced mass. It is important to recognize that although $\mathcal{V}$ is very small, the monopole term should not be taken to be zero because the dipole contribution might also be very small. If the scattering body is homogeneous with mass density $\rho_{\mathrm{b}}$, then we can set $\mathcal{V}=M / \rho_{b}$.

EXAMPLE 12.2 The sketch shows an arrangement in which projector $A$ and hydrophone $B$ are situated along orthogonal lines relative to the center of a scattering body that is rigid and stationary. The distance from the source to the body is sufficiently large to consider the incident wave to be planar, and the wavelength of the incident wave is much greater than the body's largest dimension. The body is axisymmetric relative to the $z$-axis, which is coplanar with $\bar{e}_{\mathrm{I}}$ and $\bar{e}_{r}$. The virtual mass matrix is described relative to the volume of the body, according to $W_{1,1}=W_{2,2}=5 \mathcal{V}, \Lambda_{3,3}=\mathcal{V}$. Determine the dependence on the angle of incidence $\psi$ of the scattered pressure at point $B$, normalized according to $4 \pi r\left|P_{\mathrm{S}}\left(\bar{x}_{A}\right) / P_{\mathrm{I}}\right| /\left(k^{2} \mathcal{V}\right)$.


Figure 1.

## Significance

Considering a specific case of scattering from a rigid body gives a clearer picture of the angular dependence, as well as some understanding to the influence of the virtual added mass.

## Solution

To employ Eq. (12.3.22), we must define the direction angles. We define the $x$-axis to lie in the plane formed by $\bar{e}_{\mathrm{I}}$ and $\bar{e}_{r}$, which also contains the $z$-axis. This is allowable because it does not matter how much the body is rotated about its axis of symmetry. The formula was derived on the basis that $\bar{e}_{\mathrm{I}}$ has a positive component in the $x$ direction, which means that the $x$-axis is down and to the right in the sketch. Thus, the unit vectors are

$$
\begin{equation*}
\bar{e}_{\mathrm{I}}=\cos \psi \overline{\boldsymbol{e}}_{z}+\sin \psi \overline{\boldsymbol{e}}_{x}, \quad \bar{e}_{r}=\sin \psi \overline{\boldsymbol{e}}_{z}-\cos \psi \overline{\boldsymbol{e}}_{x} \tag{1}
\end{equation*}
$$

A comparison of these expressions to the general description in Eq. (12.3.21) leads to recognition that $\psi_{\mathrm{I}}=\psi, \theta_{\mathrm{I}}=0, \psi_{r}=\pi / 2-\psi$, and $\theta_{r}=\pi$. The result is that

$$
\begin{equation*}
\frac{4 \pi r}{k^{2} \mathcal{V}}\left|P_{\mathrm{S}}\left(\bar{x}_{0}\right)\right|=1+\frac{1}{\mathcal{V}}\left(W_{1,1}-W_{3,3}\right) \cos \psi \sin \psi=1+\frac{\left(W_{1,1}-W_{3,3}\right)}{2 \mathcal{V}} \sin (2 \psi) \tag{2}
\end{equation*}
$$

This function is depicted in Fig. 2. Note that this is not a directivity. Rather, a radial line at a specific $\psi$ indicates the magnitude of $\left|P_{s}\right|$ when the axis of symmetry, which is $x$, is aligned at that angle.


Figure 2.

The plot has interesting features, some of which might be considered to be counterintuitive. At $\psi=0,90^{\circ}, 180^{\circ}$, and $270^{\circ}$, the normalized scattered pressure is one. This is the monopole effect. But it is one also at $\psi=135^{\circ}$ and $\psi=315^{\circ}$. At these angles, the dipole contribution associated with the amplitude of the dipole contribution is twice the monopole value, but it is $180^{\circ}$ out-of-phase from the monopole. In contrast, the scattered pressure arriving at hydrophone $B$ has a maximum at $\psi=45^{\circ}$ and $\psi=225^{\circ}$ because the two contributions are in-phase. The nulls which occur at $\psi=105^{\circ}, 165^{\circ}, 285^{\circ}$, and $245^{\circ}$ result from the two effects having equal magnitude by opposite phase.

Many individuals think of scattering as a mirror-like effect, which is the viewpoint of geometrical acoustics. In such thinking, $\psi=45^{\circ}$ and $\psi=225^{\circ}$ are like placing a mirror in the $x y$-plane. The angle of incidence of $\bar{e}_{\text {I }}$ on this plane equals the angle of reflection of $\bar{e}_{r}$, which should lead to a maximum pressure at hydrophone $B$. This expectation is confirmed by Fig. 1, but that does not mean that the underlying reasoning is correct. If it were, then $\psi=135^{\circ}$ and $\psi=270^{\circ}$ would be nulls because the corresponding angle of reflection is downward. In general, low-frequency scattering is not a process of reflection.

## Effect of Rigid Body Motion

We now turn to the fact that the scattering body often is immersed in the fluid without support, as in the case of a rain drop. Such movement generates a radiating pressure field that adds to the pressure field obtained when the body is stationary. Section 7.4.2 developed a multipole expansion that describes low-frequency radiation from an oscillating rigid body. That development led to the observation that rotation of a body at low frequencies generates a very weak quadrupole that is negligible if
the dipole contribution is nonzero. This is the present situation, so we only need to consider translational motion, which is governed by Newton's second law.

The most expedient way to proceed is to return to Eq. (12.1.3), which is the velocity continuity condition for any scattering problem,

$$
\begin{equation*}
\bar{n} \cdot \nabla P_{\mathrm{s}}=-\bar{n} \cdot \nabla P_{\mathrm{I}}-i \omega \rho_{0} \bar{n} \cdot \bar{V}, \quad \bar{x} \in \mathcal{S} \tag{12.3.23}
\end{equation*}
$$

This relation allows us to regard the scattered field as the superposition from two incident plane waves: the actual one, and one whose particle velocity is $-\bar{V}$. The pressure amplitude of the latter is $\rho_{0} c V$. In other words, the effective incident pressure is

$$
\begin{equation*}
\left(P_{\mathrm{I}}\right)_{\text {effective }}=\hat{P}_{\mathrm{I}} e^{-i k \bar{e}_{1} \cdot \bar{x}}-\rho_{0} c V e^{-i k \bar{e}_{v} \cdot \bar{x}} \tag{12.3.24}
\end{equation*}
$$

The first place we introduce this replacement is the analysis of the monopole part of the scattered pressure, Eq. (12.3.17). This effect does not depend on the direction in which the incident wave propagates, but the phase lag between the incident velocity and $\bar{V}$ is $180^{\circ}$. Therefore, the monopole amplitude is

$$
\begin{equation*}
A=-\frac{k^{2} \mathcal{V}}{4 \pi} A=-\frac{k^{2} \mathcal{V}}{4 \pi}\left(\hat{P}_{\mathrm{I}}-\rho_{0} c|V|\right) \tag{12.3.25}
\end{equation*}
$$

The modification of the dipole moment entails replacement of $P_{\mathrm{I}} \bar{e}_{\mathrm{I}}$ with $P_{\mathrm{I}} \bar{e}_{\mathrm{I}}-\rho_{0} c \bar{V}$. Doing so converts the dipole moment in Eq. (12.3.18) to

$$
\begin{equation*}
\{D\}=\frac{i k}{4 \pi}[\mathcal{V}[I]+[W]]\left\{\hat{P}_{\mathrm{I}}\left\{e_{\mathrm{I}}\right\}-\rho_{0} c\{V\}\right\} \tag{12.3.26}
\end{equation*}
$$

The last step in the analysis is determination of $\bar{V}$. With that as our objective, we formulate Newton's Second Law for the object. It equates the inertia of the body to the resultants of the incident and scattered pressure distributions. Let $M$ be the body's mass. The complex amplitude of its acceleration is $i \omega \bar{V}$, so it must be that

$$
\begin{equation*}
M(i \omega \bar{V})=\bar{F}_{\mathrm{I}}+\bar{F} \tag{12.3.27}
\end{equation*}
$$

The scattered pressure resultant is described by (12.3.20), with the incident pressure set to $P_{\mathrm{I}} \overline{\mathrm{e}}_{\mathrm{I}}-\rho_{0} c \bar{V}$,

$$
\begin{equation*}
\{F\}=i k[W]\left\{\hat{P}_{\mathrm{I}}\left\{e_{\mathrm{I}}\right\}-\rho_{0} c\{V\}\right\} \tag{12.3.28}
\end{equation*}
$$

The resultant of the incident pressure is

$$
\begin{equation*}
\bar{F}_{\mathrm{I}}=\iint_{\mathcal{S}}\left(-P_{\mathrm{I}} \bar{n}\right) d \mathcal{S} \tag{12.3.29}
\end{equation*}
$$

The sense of $\bar{n}$ is outward from the region $\mathcal{V}$ enclosed by $\mathcal{S}$, so the divergence theorem converts this to

$$
\begin{equation*}
\bar{F}_{\mathrm{I}}=-\iiint_{\mathcal{V}}\left(\nabla P_{\mathrm{I}}\right) d \mathcal{V} \tag{12.3.30}
\end{equation*}
$$

The scale over which $P_{\mathrm{I}}$ varies is an acoustic wavelength, which is restricted here to being much larger than the size of the body. Therefore, we may consider $\nabla P_{\mathrm{I}}$ to be a constant in this expression. For a plane wave, $\nabla P_{\mathrm{I}}=-i k \hat{P}_{\mathrm{I}} \bar{e}_{\mathrm{I}} \exp \left(-i k \bar{e}_{\mathrm{I}} \cdot \bar{x}\right)$. We evaluate this gradient at the origin and invoke the smallness of $k \bar{x}$, so that

$$
\begin{equation*}
\bar{F}_{\mathrm{I}}=i k \mathcal{V} \hat{\mathcal{P}}_{\mathrm{I}} \bar{e}_{\mathrm{I}} \Longleftrightarrow\left\{F_{\mathrm{I}}\right\}=i k \mathcal{V} \hat{P}_{\mathrm{I}}\left\{\bar{e}_{\mathrm{I}}\right\} \tag{12.3.31}
\end{equation*}
$$

Substitution of $\{F\}$ and $\left\{F_{\mathrm{I}}\right\}$ into Newton's law gives

$$
\begin{equation*}
\operatorname{Mi\omega }\{V\}=i k[W]\left\{\hat{P}_{\mathrm{I}}\left\{e_{\mathrm{I}}\right\}-\rho_{0} c\{V\}\right\}+i k \mathcal{V} \hat{P}_{\mathrm{I}}\left\{\bar{e}_{\mathrm{I}}\right\} \tag{12.3.32}
\end{equation*}
$$

A rearrangement of terms yields an expression that may be solved for $\{V\}$, specifically

$$
\begin{equation*}
\left[\frac{M}{\rho_{0}}[I]+[W]\right]\{V\}=\frac{\hat{P}_{\mathrm{I}}}{\rho_{0} c}[\mathcal{V}[I]+[W]]\left\{\bar{e}_{\mathrm{I}}\right\} \tag{12.3.33}
\end{equation*}
$$

This expression confirms that it was correct to say in Chap. 7 that $\rho_{0}[W]$ is an added mass.

The closure of the analysis uses the value of $\{V\}$ found from the preceding to form the dipole moment in Eq. (12.3.26). Then, the expressions for the monopole amplitude and dipole moment are substituted into the mulipole expansion, Eq. (7.4.23). Letting $G\left(\bar{x}_{0}, \bar{x}_{C}\right)$ be the free-space Greens function ultimately leads to

$$
\begin{align*}
P_{\mathrm{S}}\left(\bar{x}_{0}\right)= & -\frac{k^{2}}{4 \pi}\left[\mathcal{V}\left(\hat{P}_{\mathrm{I}}-\rho_{0} c|V|\right)-\left\{e_{r}\right\}^{\mathrm{T}}[\mathcal{V}[I]\right. \\
& \left.+[W]]\left\{\hat{P}_{\mathrm{I}}\left\{e_{\mathrm{I}}\right\}-\rho_{0} c\{V\}\right\}\left(1-\frac{i}{k r}\right)\right] \frac{e^{-i k r}}{r} \tag{12.3.34}
\end{align*}
$$

As is true for scattering from a stationary rigid body, a quantitative evaluation of this expression requires that we know [ W ]. The values for a sphere and thin circular disk are provided in Eq. (7.4.38) and at the concluding paragraph of the preceding section.

The special case of a neutrally buoyant object, such as a submarine, is interesting. The mass is the same as that of the displaced fluid, so $M / \rho_{0}=\mathcal{V}$. According to Eq. (12.3.33), this situation leads to $\rho_{0} c \bar{V}=\hat{P}_{\mathrm{I}} \bar{e}_{\mathrm{I}}$, from which it follows that $P_{\mathrm{s}}=0$. This situation occurs because the neutral buoyant body moves in unison with the fluid, so it does not impede the propagation of the incident wave. This phenomenon assumes that the acoustic wavelength is much larger than any dimension of the object.

It also ignores the fact that no object is truly rigid. The role of structural flexibility is examined in the last section of this chapter.

EXAMPLE 12.3 A thin aluminum disk is suspended in a large vat of water in a manner that allows it to be rotated about its vertical diameter without otherwise restraining the disk. A plane wave traveling horizontally is incident on the disk, with the angle between $-\bar{e}_{\mathrm{I}}$ and the normal to the disk being $\psi_{\mathrm{I}}$. The mass of the disk is 1.7 g , and its radius is 20 mm . The frequency of the incident wave is 100 Hz , and its amplitude is 10 kPa . (a) Determine the magnitude and direction of the disk's complex velocity relative to the $z$ direction as a function of $\psi_{\mathrm{I}}$. (b) Determine the amplitude of the backscattered pressure as a function of $\psi_{\mathrm{I}}$ and compare the result to backscatter for the case of a stationary disk.

## Significance

The focus is on the tensorial nature of the added mass effect and how it affects the backscattered pressure. The quantitative results illuminate the general effect of allowing an object to move.

## Solution

The coordinate system and geometry for this system are shown in Fig. 1. The given parameters correspond to $k a=8.491\left(10^{-3}\right)$, which certainly is sufficiently small to warrant application of Rayleigh scattering theory. We begin by assembling the basic quantities. According to Eq. (7.4.38), the virtual mass tensor is a single element, $\Lambda_{3,3}=(8 / 3) a^{3}$. It is not surprising that only the 3,3 element is nonzero, because incidence of a wave traveling parallel to the surface of the disk will not lead to significant scattering. Another consideration is that $\mathcal{V}$ is not given. We could take it be zero, based on thinness of the disk, but the monopole contribution is dominant at orientations for which the dipole contribution is very small. Therefore, we determine it from the mass, which gives $\mathcal{V}=M / \rho_{\mathrm{a}}$.


Figure 1

Axisymmetry allows us to situate $\bar{e}_{\mathrm{I}}$ in the $x z$-plane, so

$$
\left\{e_{\mathrm{I}}\right\}=\left[\begin{array}{lll}
\sin \psi_{\mathrm{I}} & 0 & -\cos \psi_{\mathrm{I}}
\end{array}\right]^{\mathrm{T}}
$$

All terms in Eq. (12.3.33) have been characterized, so we may solve that equation for the complex velocity amplitude. Doing so leads to

$$
\{V\}=\frac{\hat{P}_{\mathrm{I}}}{\rho_{0} c}\left[\left[\begin{array}{ccc}
\frac{M}{\rho_{0}} & 0 & 0  \tag{1}\\
0 & \frac{M}{\rho_{0}} & 0 \\
0 & 0 & \frac{M}{\rho_{0}}+\frac{8}{3} a^{3}
\end{array}\right]\right]^{-1}\left[\begin{array}{ccc}
\frac{M}{\rho_{\mathrm{al}}} & 0 & 0 \\
0 & \frac{M}{\rho_{\mathrm{al}}} & 0 \\
0 & 0 & \frac{M}{\rho_{\mathrm{al}}}+\frac{8}{3} a^{3}
\end{array}\right]\left\{\begin{array}{c}
\sin \psi_{\mathrm{I}} \\
0 \\
-\cos \psi_{\mathrm{I}}
\end{array}\right\}
$$

The matrix being inverted is diagonal, so we find that

$$
\begin{gather*}
V_{x}=\left(\frac{\rho_{0}}{\rho_{\mathrm{al}}}\right) \frac{\hat{P}_{\mathrm{I}}}{\rho_{0} c} \sin \psi_{\mathrm{I}}, \quad V_{y}=0  \tag{2}\\
V_{z}=-\left(\frac{M+(8 / 3) \rho_{\mathrm{al}} a^{3}}{M+(8 / 3) \rho_{0} a^{3}}\right)\left(\frac{\rho_{0}}{\rho_{\mathrm{al}}}\right) \frac{\hat{P}_{\mathrm{I}}}{\rho_{0} c} \cos \psi_{\mathrm{I}}
\end{gather*}
$$

Both complex velocity components are real, so their relative phase is zero. This corresponds to a temporal velocity vector that is oriented in a constant direction. It lies in the $x z$-plane, so we may describe $\bar{V}$ in terms of its magnitude $|V|$ and angle $\theta$ relative to the $z$-axis as $V_{x}=|\bar{V}| \sin \theta, V_{z}=|\bar{V}| \cos \theta$. (Both components are required to assure that $\theta$ is assigned to the proper quadrant.) The velocity components, as well as $|V|$ and $\theta$, are graphed in Fig. 2. Perhaps the most unexpected feature is the fact that $\bar{V}$ is parallel to $\bar{e}_{z}$ when $\psi_{\mathrm{I}}=90^{\circ}$. Although the disk is thin, it moves at this incidence because its volume is finite.


Figure 2

The complex velocity amplitude is used to compute $(r / a)\left|P_{\mathrm{I}}\right| /\left(\rho_{0} c\right)$ according to Eq. (12.3.34). We are interested in the farfield backscatter, so we set $\bar{e}_{r}=-\bar{e}_{\mathrm{I}}$ and drop the term that is the reciprocal of $k r$. Thus, we have

$$
\begin{align*}
\left(\frac{r}{a}\right) \frac{\left|P_{s}\left(\bar{x}_{0}\right)\right|}{\rho_{0} c}= & -\frac{(k a)^{2}}{4 \pi}\left(\frac{M}{\rho_{\mathrm{al}} a^{2}}\right)\left(\hat{P}_{\mathrm{I}}-\rho_{0} c|V|\right) \\
& +\frac{(k a)^{2}}{4 \pi}\left\{\begin{array}{c}
\sin \psi_{\mathrm{I}} \\
0 \\
-\cos \psi_{\mathrm{I}}
\end{array}\right\}\left[\begin{array}{ccc}
\frac{\mathrm{T}}{\frac{\mathrm{\rho}}{\mathrm{al}}} & 0 & 0 \\
0 & \frac{M}{\rho_{\mathrm{al}}} & 0 \\
0 & 0 & \frac{M}{\rho_{\mathrm{al}}}+\frac{8}{3} a^{3}
\end{array}\right]  \tag{3}\\
& \times\left\{\hat{P}_{\mathrm{I}}\left\{\begin{array}{c}
-\sin \psi_{\mathrm{I}} \\
0 \\
\cos \psi_{\mathrm{I}}
\end{array}\right\}-\rho_{0} c\left\{\begin{array}{c}
V_{x} \\
0 \\
V_{z}
\end{array}\right\}\right\}
\end{align*}
$$

Figure 3 is constructed by substituting $\{V\}$ at each $\psi_{I}$ into Eq. (3). The value of $\left|P_{s}\right|$ for the case of a stationary disk is obtained by setting $\{V\}=\{0\}$.


Figure 3

It can be seen that the backscatter is much greater if the disk is immobilized. An explanation of this behavior may be found in an evaluation of the ratio of $V_{x}$ and $V_{z}$ to the respective components of $\bar{V}_{\mathrm{I}}$, which are $V_{\mathrm{I}} \sin \psi_{\mathrm{I}}$ and $-V_{\mathrm{I}} \cos \psi_{\mathrm{I}}$. Equation (2) indicate that

$$
\begin{equation*}
\frac{V_{x}}{\bar{V}_{\mathrm{I}} \cdot \bar{e}_{x}}=\frac{\rho_{0}}{\rho_{\mathrm{al}}}=0.370, \frac{V_{z}}{\bar{V}_{\mathrm{I}} \cdot \bar{e}_{z}}=\left(\frac{M+(8 / 3) \rho_{\mathrm{al}} a^{3}}{M+(8 / 3) \rho_{0} a^{3}}\right)\left(\frac{\rho_{0}}{\rho_{\mathrm{al}}}\right)=0.954 \tag{12.3.35}
\end{equation*}
$$

Movement of the disk in the $\bar{e}_{x}$ direction leads to little scattering, so the mismatch between $V_{x}$ and $\bar{V}_{I} \cdot \bar{e}_{x}$ has little effect on the scattering. In contrast, movement of the disk in the $\bar{e}_{z}$ direction is important. This velocity component is very nearly equal to $\bar{V}_{\mathrm{I}} \cdot \bar{e}_{z}$. If they truly were equal, there would be no scattering. This phenomenon is comparable to ocean waves. If they are incident on a stationary object, there is a large splash. In contrast, a floating object induces little ripple. Another aspect to note is that the scattered pressure is essentially zero at $\psi_{\mathrm{I}}=90^{\circ}$ and $270^{\circ}$. At these angles, the contribution associated with $[W$ ] is zero. The remaining terms are negligible because $M / \rho_{\mathrm{al}}=\mathcal{V}$ is much less than $a^{2}$.

### 12.4 Measurements and Metrics

Acoustic scattering is the core phenomenon for a number of applications, ranging from biomedical ultrasonics to fishing to naval sonar. Thus, it is undoubtedly true that scattered signals are measured much more frequently than they are evaluated analytically. This bias is even stronger if we include marine mammals and bats, who have remarkable capabilities to "see" with sound.

Sonar has two implementations. Passive sonar is a process of listening, in that it detects sound radiated by a body. If the object that is to be located is stationary and does not vibrate, it will emit no sound and therefore cannot be located. Active sonar does not suffer from this shortcoming. It is used recreationally and commercially to locate fish. Typically, this entails broadcasting a signal from the boat, then measuring the signal returned to the boat. Suitable processing of the returned signal can determine whether it is scattered by fish, and if so, where the fish are located. This application is a monostatic scattering measurement. The "mono" prefix is used because signal generation and measurement are at a single location. Many of the sonar concepts follow those for radar, wherein "static" refers to the apparatus being stationary. Monostatic scattering is the primary concept for biomedical ultrasound, wherein the device that generates the probing signal also detects the signal scattered by a heterogeneous region, such as a tumor.

Naval sonar systems applications rely on both monostatic and bistatic scattering measurements. In a bistatic measurement, the incident signal is generated at one location and the scattered signal is detected at another. This can be done from an airplane or helicopter by dropping sonobuoys, or it can be done from ships and submarines by using a towed array.

Ultrasonic imaging for medical diagnostics is the acoustical analog of our visual sense, in that they rely on contrast. The field scattered by a heterogeneity is scanned spatially. Strong returns correspond to greater impedance mismatch, so they are displayed as lighter regions. Weaker returns correspond to little or no heterogeneity, and therefore are displayed as dark regions. The field properties of the scattered field, such as its multipole resolution, typically is not an issue in biomedical applications.

The underlying properties of the scattered field are of vital importance to sonar applications. The general objective there is to deduce the nature of the scattering body from measurements. We shall not delve into how the measured signal is used to identify the fundamental nature of the scattering body, which is a process of target identification. However, if we consider the result for the field scattered by a rigid body, we can gain a hint at the possibility of using scattering data to characterize the scatterer. The virtual mass matrix $\rho_{o}[W]$ in Eq. (12.3.19) is symmetric, so it contains six unknown elements. Another unknown parameter is the volume of the scattering body. Thus, if we have several measurements, corresponding to angles of incidence and/or different measurement locations, it is conceivable that we could infer these unknown parameters.

Obviously, this notion relies on several assumptions. Several sources and/or receivers are required to obtain bistatic measurements. It is necessary to determine
the location of the body. Range information can be determined by using a pulse, rather than a continuous wave. The simplest measurement uses the time $\tau$ for a signal to return to the source location in a monostatic measurement. Because the travel time to the body is $\tau / 2$, the range will be $c \tau / 2$. Orientation of the line from the source/receiver location to the body can be determined by rotating a highly directive source until the scattered signal is largest, or by using a phased array. Another assumption is that the orientation of the scattering body does not change between measurements. This can be addressed by taking the measurements over a short interval. The most fundamental assumption buried in the notion of identifying $\mathcal{V}$ and $[\Lambda]$ is that the analysis is valid. If the body is flexible, then the scattered field will be much more complicated than a simple superposition of monopole and dipole fields. Indeed, in that case, the nature of the field is likely to be highly dependent on the frequency. An efficient way of performing scattering measurements over a frequency range is to alter the pulse from a pure tone burst. The shape of the envelope can be adjusted (amplitude modulation), and the frequency can be varied (frequency modulation) within the pulse's interval. Then, FFT processing of the scattered signals, which also will be pulses, will yield frequency spectra that can be converted to transfer functions.

Unless the duration of the pulse is very long relative to the range divided by the sound speed, the scattered signal will be measured distinctly from the incident. The logical question is what to do with it? Other than travel time and location, the most fundamental metric is based on a concept developed for radar. It originates from a hypothetical situation. Suppose a planar harmonic wave is normally incident on a rigid stationary disk. Let us further suppose that the planar wave is a confined beam whose cross section exactly matches the disk. This is the situation depicted in Fig. 12.3. In the high-frequency limit of ray acoustics, the scattered signal $P_{\mathrm{s}}$ is the regular reflection of the incident signal, corresponding to a reflection coefficient of one. In that case, the power transported in the reflected signal is the same as that of the incident signal. The magnitude of the time-averaged intensity in the incident wave is $I_{\mathrm{I}}=\left|P_{\mathrm{I}}\right|^{2} /\left(2 \rho_{0} c\right)$, so the power in the scattered signal is $\left(\mathcal{P}_{\mathrm{s}}\right)_{\mathrm{av}}=I_{\mathrm{I}} \mathcal{A}$, where $\mathcal{A}$ is the cross-sectional area of the side of the disk that is irradiated by the incident signal. Division of the time-averaged power in the reflected wave by the intensity of the incident wave yields a measure of the cross-sectional area that blocks the incident wave.

Fig. 12.3 Model of scattering from a disk used to define the scattering cross section. The picture generally is not descriptive of the actual scattering process


To extend this idea to the actual field scattered by an arbitrary body, the timeaveraged power transported in the scattered signal is computed by integrating over a sphere $\mathcal{S}_{0}$ that surrounds the scatterer. The spherical coordinates of a point on the sphere's surface are $r_{0}, \psi_{0}, \theta_{0}$. The radius $r_{0}$ of this sphere is taken to be much greater than $2 \pi / k$, as well as the largest dimension of the body. Then, the radial velocity on the sphere's surface is $V_{r}=P_{\mathrm{s}}\left(r_{0}, \psi_{0}, \theta_{0}\right) /\left(\rho_{0} c\right)$. The general definition of the scattering cross section is that it is the time-averaged power transported by the scattered signal divided by the intensity of the incident wave,

$$
\begin{equation*}
\sigma_{\mathrm{s}}=\frac{\left(\mathcal{P}_{\mathrm{s}}\right)_{\mathrm{av}}}{I_{\mathrm{I}}}=\iint_{\mathcal{S}_{0}} \frac{I_{\mathrm{s}}\left(r_{0}, \psi_{0}, \theta_{0}\right)}{I_{\mathrm{I}}} d \mathcal{S}_{0}, \quad I_{\mathrm{s}}=\frac{\left|P_{\mathrm{s}}\right|^{2}}{2 \rho_{0} c} \tag{12.4.1}
\end{equation*}
$$

The intensity in the farfield is proportional to $1 / r_{0}^{2}$, so setting $d \mathcal{S}_{0}=r_{0}^{2}\left(\sin \psi_{0}\right)$ $d \theta_{0} d \psi_{0}$ will yield a value for $\sigma_{\mathrm{s}}$ that is independent of $r_{0}$.

In the hypothetical model of scattering from a disk, $\sigma_{\mathrm{s}}=\mathcal{A}$. This leads to interpretation of $\sigma_{\mathrm{s}}$ as the apparent area of the scattering object that blocks the incident wave. Because the field scattered by an arbitrary body depends on the orientation of $\bar{e}_{\text {I }}$ relative to the body, the scattering cross section depends on that orientation. The value of $\sigma_{\mathrm{s}}$ might exceed the area of the body's projection on to the plane normal to the $\bar{e}_{\mathrm{I}}$, or it might be much smaller.

Consider the case of scattering from a stationary rigid body. Equation (12.3.19) describes the scattered field as the sum of a monopole whose strength is proportional to the displaced mass $\rho_{0} \mathcal{V}$, and a dipole whose moment is proportional to $[\mathcal{V}[I]+[W]]\left\{e_{\mathrm{I}}\right\}$. It is useful to identify the direction of the dipole moment corresponding to a known $[W]$ because scattering associated with the dipole will be strongest in that and the opposite direction. Toward that objective, we write the vectorial part of the dipole moment in terms of its magnitude $\Phi$ and a unit vector $\left\{e_{D}\right\}$ that is the direction of the vector. In other words, we set $\Phi\left\{e_{D}\right\}=[\mathcal{V}[I]+[W]]\left\{e_{\mathrm{I}}\right\}$. It follows that

$$
\begin{equation*}
\Phi=\left|[\mathcal{V}[I]+[W]]\left\{e_{\mathrm{I}}\right\}\right|, \quad\left\{e_{D}\right\}=\frac{[\mathcal{V}[I]+[W]]\left\{e_{\mathrm{I}}\right\}}{\Phi} \tag{12.4.2}
\end{equation*}
$$

This transformation yields a description of $\left\{e_{D}\right\}$ in terms of components relative to the $x y z$ coordinate system associated with the known components of $[W]$. We also represent the direction $\left\{e_{r}\right\}$ from the origin to a field point on the surface of the surrounding sphere $\mathcal{S}_{0}$ in terms of components relative to the same coordinate system. Doing so allows us to evaluate product $\left\{e_{r}\right\}^{\mathrm{T}}[[\mathcal{V}[I]+[W]]]\left\{e_{\mathrm{I}}\right\}$, which now is $\Phi\left\{e_{r}\right\}^{\mathrm{T}}\left\{e_{D}\right\}$. In other words, the directivity of the dipole is the dot product of $\bar{e}_{r}$ and $\bar{e}_{D}$, and the strength of the dipole is proportional to $\Phi$. Let $\psi_{0}$ be the angle between the vectors, that is,

$$
\begin{equation*}
\psi_{0}=\cos ^{-1}\left(\left\{e_{r}\right\}^{\mathrm{T}}\left\{e_{D}\right\}\right) \tag{12.4.3}
\end{equation*}
$$

This corresponding representation of the farfield pressure scattered by a stationary rigid body is

$$
\begin{equation*}
P_{\mathrm{S}}\left(\bar{x}_{0}\right)=-\frac{k^{2} \hat{P}_{\mathrm{I}}}{4 \pi}\left[\mathcal{V}-\Phi \cos \psi_{0}\right] \frac{e^{-i k r_{0}}}{r_{0}} \tag{12.4.4}
\end{equation*}
$$

Integration of this expression over the surface of the sphere yields the radiated power. We define the polar direction for a set of spherical coordinates to be $\bar{e}_{D}$. The field is axisymmetric with respect to an axis aligned with $\left\{e_{D}\right\}$, so the differential area element is set to $2 \pi r_{0}^{2} \sin \psi_{0} d \psi_{0}$. Division of $\mathcal{P}_{\mathrm{s}}$ by $\left|I_{\mathrm{I}}\right|=\left|\hat{P}_{\mathrm{I}}\right|^{2} /\left(2 \rho_{0} c\right)$ yields the scattering cross section. The result is

$$
\begin{equation*}
\sigma_{\mathrm{s}}=\frac{k^{4}}{4 \pi}\left(\mathcal{V}^{2}+\frac{1}{3}\left|[\Phi]\left\{e_{\mathrm{I}}\right\}\right|^{2}\right) \tag{12.4.5}
\end{equation*}
$$

It should be noted that the integration over $\mathcal{S}_{0}$ has removed any dependence on the orientation of the dipole moment, so there is no need to actually evaluate $\left\{e_{D}\right\}$. Specific results are

$$
\begin{gather*}
\text { Sphere: } \sigma_{\mathrm{s}}=\frac{7}{9}\left(\pi a^{2}\right)(k a)^{4}  \tag{12.4.6}\\
\text { Thin circular disk: } \sigma_{\mathrm{s}}=\frac{16}{27 \pi^{2}}\left(\pi a^{2}\right)(k a)^{4}\left(\cos \psi_{\mathrm{I}}\right)^{2}
\end{gather*}
$$

where $\psi_{\mathrm{I}}$ is the angle between the normal to the disk and the incident wave's direction.
In the case of a sphere, the factor $\pi a^{2}$ is the physical area that blocks the incident wave at any angle. For a thin disk, the maximum blocking area is $\pi a^{2}$, which corresponds to normal incidence. At arbitrary incidence, the physical blocking area is $\pi a^{2} \cos \psi_{\mathrm{I}}$, so the above expression indicates that the decrease of $\sigma$ when $\psi_{\mathrm{I}}$ is increased is due to more than the decrease of the blocking area. The Rayleigh scattering limit describes cases where $k a \ll 1$, so the dependence of $\sigma_{\mathrm{s}}$ for each shape on $(k a)^{4}$ tells us that the bodies appear to be much smaller than their cross-sectional area.

The scattering cross section represents a spatial average of what is seen at various field points. A different metric is used to convey what is seen at a specific location. This metric is defined in terms of a solid angle, which is a geometric parameter we encountered in Sect.7.4.1. A solid angle will arise in the next section, so it is helpful to review the concept. Consider the normal $\bar{n}\left(\bar{x}_{\mathrm{s}}\right)$ to the surface. The intersection of the surface and any plane that contains this normal is a curve. We are interested in the differential arc of this curve that is centered at $\bar{x}_{\mathrm{s}}$. The center of curvature of this arc is the intersection of $\bar{n}\left(\bar{x}_{\mathrm{s}}\right)$ and any other normal to the arc, such as that at either end. The distance from the center of curvature to the surface is the radius of curvature of the arc. The plane containing $\bar{n}\left(\bar{x}_{\mathrm{s}}\right)$ is defined by its rotation about $\bar{n}\left(\bar{x}_{\mathrm{s}}\right)$. At some rotation angle, the arc has the largest radius of curvature. A further rotation by $90^{\circ}$
forms an arc with the smallest radius. These planes and arcs are depicted in Fig. 12.4. The radii $r_{1}$ and $r_{2}$ are the principal radii of curvature. (For a sphere $r_{1}=r_{2}=a$, whereas $r_{1}=a$ and $r_{2}$ is infinite for a cylinder.) The arclengths are $d \ell_{1}=r_{1} d \theta_{1}$ and $d \ell_{2}=r_{1} d \theta_{2}$ where $d \theta_{n}$ is the angle between the normals to the surface at the ends of $d \ell_{n}$. The area of the patch that is formed is $d \mathcal{S}=d \ell_{1} d \ell_{2}$. The definition of a solid angle is

$$
\begin{equation*}
d \Gamma=\frac{d \mathcal{S}}{r_{1} r_{2}} \tag{12.4.7}
\end{equation*}
$$

where $\Gamma$ is the symbol we shall employ for the solid angle. Thus, the solid angle subtended by the patch is $d \Gamma=d \theta_{1} d \theta_{2}$.

Fig. 12.4 Surface patch formed from arcs whose radii of curvature are the principal values at a point


The scattering cross section is defined in Eq. (12.4.1) as an integral over a sphere. Both principal radii of that surface are $r_{0}$, so the differential element of area may be written as $d \mathcal{S}_{0}=r_{0}^{2} d \Gamma$. Concurrently, any integral may be represented as the antiderivative of its integrand. Thus, the definition of $\sigma_{\mathrm{s}}$ may be written in two ways, as

$$
\begin{equation*}
\sigma_{\mathrm{s}}=\iint_{\mathcal{S}_{0}} \frac{I_{\mathrm{s}}\left(r_{0}, \psi_{0}, \theta_{0}\right)}{I_{\mathrm{I}}}\left(r_{0}^{2} d \Gamma\right)=\iint_{\mathcal{S}_{0}} \frac{d \sigma_{\mathrm{s}}}{d \Gamma} d \Gamma \tag{12.4.8}
\end{equation*}
$$

The quantity $d \sigma_{\mathrm{s}} / d \Gamma$ is the differential scattering cross section. Matching the above integrands gives

$$
\begin{equation*}
\frac{d \sigma_{\mathrm{s}}}{d \Gamma}=\frac{I_{\mathrm{s}}\left(r_{0}, \psi_{0}, \theta_{0}\right)}{I_{\mathrm{I}}} r_{0}^{2}=\frac{r_{0}^{2}\left|P_{\mathrm{s}}\left(r_{0}, \psi_{0}, \theta_{0}\right)\right|^{2}}{\left|P_{\mathrm{I}}\right|^{2}} \tag{12.4.9}
\end{equation*}
$$

In other words, $d \sigma_{\mathrm{s}} / d \Gamma$ is proportional to the directivity of the farfield intensity. It is implicit to this expression that the scattering cross section is defined for a specific orientation of $\bar{e}_{\mathrm{I}}$ relative to the body, so that $\psi_{0}$ and $\theta_{0}$ are defined relative to a bodyfixed direction. If we define that direction to be $\bar{e}_{r}$, then backscatter corresponds to $\bar{e}_{r}=-\bar{e}_{\mathrm{I}}\left(\psi_{0}=\pi\right)$. A measurement at an arbitrary field point gives the differential scattering cross section in a bistatic measurement. The special case where $\bar{e}_{r}=\bar{e}_{\mathrm{I}}$ ( $\psi_{0}=\pi$ ) is said to be forward scatter.

A quantity that causes confusion for some is the backscatter cross section, which is denoted as $\sigma_{\text {back }}$. Despite the name, it is not derived from the actual scattering cross section. Rather, it is the value of the scattering cross section that would be obtained if the scattering was isotropic at the backscatter value. Accordingly, if $d \sigma_{\mathrm{s}} / d \Gamma$ were constant at its value for $\bar{e}_{r}=-\bar{e}_{\mathrm{I}}$, integrating it over $\mathcal{S}_{0}$ would give a $4 \pi$ factor, so that

$$
\begin{equation*}
\sigma_{\text {back }}=4 \pi\left(\frac{d \sigma}{d \Gamma}\right)_{\bar{e}_{r}=-\bar{e}_{k}}=4 \pi r_{0}^{2} \frac{I_{\mathrm{s}}\left(r_{0}, \pi, 0\right)}{I_{\mathrm{I}}}=4 \pi r_{0}^{2}\left|\frac{P_{\mathrm{s}}\left(r_{0}, \pi, 0\right)}{P_{\mathrm{I}}}\right|^{2} \tag{12.4.10}
\end{equation*}
$$

Like the other metrics, $\sigma_{\text {back }}$ depends on the orientation of the incident signal relative to the body. Because $\left|P_{\mathrm{s}}\right|$ is proportional to $1 / r_{0}$, the value of $\sigma_{\text {back }}$ is independent of $r_{0}$.

Typically, the backscatter cross section is described logarithmically as the target strength. The definition is

$$
\begin{equation*}
T S=10 \log \left(\frac{\sigma_{\text {back }}}{4 \pi r_{\text {ref }}^{2}}\right) \tag{12.4.11}
\end{equation*}
$$

where $r_{\text {ref }}$ is a reference distance; the standard value is $r_{\text {ref }}=1 \mathrm{~m}$. An expression for the target strength in terms of measured pressure values results from substitution of the definition of $\sigma_{\text {back }}$ and Eq. (12.4.9) into the above. Let $\mathcal{L}_{\text {back }}$ be the sound pressure level at the source/receiver location in a monostatic measurement, and let $\mathcal{L}_{\text {I }}$ be the sound pressure level in the incident wave at that location. Then

$$
\begin{equation*}
T S=\mathcal{L}_{\text {back }}-\mathcal{L}_{\mathrm{I}}+20 \log \left(\frac{r_{0}}{r_{\text {ref }}}\right) \tag{12.4.12}
\end{equation*}
$$

Measurement of backscatter calls for a monostatic setup, in which the source and receiver are collocated, and possibly the same device. The usage of the above to evaluate target strength requires that the incident signal not be included in $\mathcal{L}_{\text {back }}$. This condition will be attained if the incident signal is a pulse that terminates before the scattered signal returns to the receiver.

### 12.5 High-frequency Approximation

Low-frequency approximations of scattering phenomena are associated with a great acoustician, Rayleigh. It is appropriate that another pioneer, Kirchhoff, should be associated with high-frequency scattering. Kirchhoff approximations are essentially applications of ray acoustics to the scattering problem. We will restrict our attention to the usual case wherein the incident wave is planar.

We begin with an analysis of backscatter from a rigid, stationary body. The first task is to locate the point on the surface at which the normal is opposite $\bar{e}_{\mathrm{I}}$. This location depends on the shape of the body and the orientation of $\bar{e}_{\mathrm{I}}$ relative to the body. This point is labeled as $C$ in Fig. 12.5, which is a side view in the first plane of principal curvature. The intensity of the backscattered wave is found by analyzing the reflection of the incident ray tube that arrives at a differential element $d \mathcal{S}$ of surface area that is centered at point $C$. The edges of $d \mathcal{S}$ are the arcs $d \ell_{1}=r_{1} d \theta_{1}$ depicted in Fig. 12.5 and $d \ell_{2}=r_{2} d \theta_{2}$ in the second plane of principal curvature. The angles $d \theta_{j}$ are the angles subtended by each relative to the respective centers of curvature.

Fig. 12.5 Description of backscatter as a process of specular reflection


The intensity of the incident plane wave is $\bar{I}_{\mathrm{I}}=\bar{e}_{\mathrm{I}}\left|P_{\mathrm{I}}\right|^{2} /\left(2 \rho_{0} c\right)$ everywhere, so the power transported by this ray tube is

$$
\begin{equation*}
d \mathcal{P}_{\mathrm{I}}=I_{\mathrm{I}} d \mathcal{S}=I_{\mathrm{I}} r_{1} r_{2} d \theta_{1} d \theta_{2} \tag{12.5.1}
\end{equation*}
$$

In the geometrical acoustics approximation, the reflection of each ray is like it would be if the surface was an infinite plane with normal $\bar{n}$. The angle of incidence at the extremities of the ray tube in Fig. 12.5 is $d \theta_{1} / 2$. The angle of reflection equals the angle of incidence, so the reflected rays are confined to a tube that is bounded by rays that are at angle $d \theta_{1}$ relative to $-\bar{e}_{\mathrm{I}}$. The consequence is that the reflected rays in this bundle represent a spherical wave that radiates from a virtual source at point $O^{\prime}$. (There is a single center for all rays because the angle of incidence for all is infinitesimal. If one were to consider incident rays at large distances from the backscatter centerline, the reflected rays would appear to radiate from centers distributed along the centerline behind point $C$. This is a manifestation of spherical aberration, which we first encountered in Example 11.1.)

The reflected rays in the plane of Fig. 12.5 are centered at point $O^{\prime}$. At a radial distance $r_{0}$, the arc subtended between the extremities of $d \ell_{1}$ is $r_{0}\left(2 d \theta_{1}\right)$. A similar description applies to the other principal plane of curvature, so the arc subtended by the rays in that plane is $r_{0}\left(2 d \theta_{2}\right)$. Thus, the reflected rays at distance $r_{0}$ cross a differential surface patch whose area is $d \mathcal{S}_{0}=4 r_{0}^{2} d \theta_{1} d \theta_{2}$. All rays contained within the scattered ray tube have the same amplitude, so the intensity along $d \mathcal{S}_{0}$ is constant at $I_{\mathrm{s}}$. The power transported across this patch in the scattered wave is

$$
\begin{equation*}
d \mathcal{P}_{\mathrm{s}}=I_{\mathrm{s}} d \mathcal{S}_{0}=I_{\mathrm{s}}\left(4 r_{0}^{2} d \theta_{1} d \theta_{2}\right) \tag{12.5.2}
\end{equation*}
$$

The surface is rigid, so the reflection coefficient for each incident ray is one. It follows that the power flow in the tube of scattered rays must equal the power flow in the tube of incident rays. Equating $d \mathcal{P}_{\mathrm{s}}$ to $d \mathcal{P}_{\mathrm{I}}$ yields

$$
\begin{equation*}
I_{\mathrm{s}}=I_{\mathrm{I}} \frac{r_{1} r_{2}}{4 r_{0}^{2}} \tag{12.5.3}
\end{equation*}
$$

The backscatter cross section is found by using this intensity to form Eq. (12.4.9), which leads to

$$
\begin{equation*}
\sigma_{\text {back }}=\pi r_{1} r_{2} \tag{12.5.4}
\end{equation*}
$$

A review of the derivation will reveal that it assumes that there is a single point $C$ on the surface at which the normal is exactly opposite to $\bar{e}_{\mathrm{I}}$. If there are multiple points that meet this criterion, the conservation of energy argument may still be applied, provided that it accounts for the contribution from each scattering site. Each contribution is a tube centered on direction $-\bar{e}_{\mathrm{I}}$. Because the scattering sites are separated by a finite distance, the tubes would not intersect. Hence, the intensity of each tube may be added as scalars to find the backscatter cross section. It also is possible that there is no point at which there is normal incidence. An example of this condition is a thin disk whose normal is not aligned with $\bar{e}_{\mathrm{I}}$. Another assumption is implicit, specifically, that the surface is convex. The situation where the surface is concave at the normal incidence point can be analyzed by merging the ray tube approach here with the analysis of focusing carried out in Example 11.1.

The preceding analysis of backscatter may be generalized to predict bistatic scattering. The first step in doing so is to locate point $C$ at which the normal to the surface is such that the angle of reflection to a designated field point equals the angle of incidence. From there, the derivation invokes concepts from differential geometry. The interested reader is referred to Pierce's presentation. ${ }^{3}$

It is much easier to evaluate the scattering cross section $\sigma_{\mathrm{s}}$ for a rigid body in the high-frequency limit. Figure 12.6 depicts all rays that are incident on the illuminated side of the body. They are confined to a (general) cylinder. The projection of the body's shape onto a plane perpendicular to $\bar{e}_{\mathrm{I}}$ defines the cross section $\mathcal{S}_{\mathrm{I}}$ of this bundle.

[^58]Fig. 12.6 Bundle of rays of a plane wave that is incident on body $\mathcal{S}$ from a cylinder whose cross section is $\mathcal{S}_{\mathrm{I}}$


The power flow in this bundle is $I_{\mathrm{I}} \mathcal{S}_{\mathrm{I}}$. Although the incident rays are reflected over a wide range of directions, their reflection is lossless. Thus, the power scattered on the illuminated side of the body equals the power contained in the incident ray tube. On the shadow side, the total pressure is essentially zero, but the total pressure is the sum of incident and scattered contributions. Thus, the power in the scattered field on that side equals the power flow in the bundle of incident rays that appear to have passed through the body. This is the same ray tube as that which is incident on the illuminated side, so the power in the scattered field on the shadow side equals the power flow in the incident side, which is $I_{\mathrm{I}} \mathcal{S}_{\mathrm{I}}$. The total power flow in the scattered field, therefore, is twice this value. Hence, we find from Eq. (12.4.9) that

$$
\begin{equation*}
\sigma_{\mathrm{s}}=2 \mathcal{S}_{\mathrm{I}} \tag{12.5.5}
\end{equation*}
$$

It is interesting to consider the geometrical acoustics results for a sphere and a thin circular disk. Both principal radii for a sphere equal the radius $a$, and the perpendicular area is $\mathcal{S}_{\mathrm{I}}=\pi a^{2}$. If the angle between the normal to the disk's surface and the incident direction is $\chi$, the cross section blocked by the disk is an ellipse whose semi-major radii are $a$ and $a \cos \chi$. Therefore, the incident ray tube area is $\mathcal{S}_{\mathrm{I}}=\pi a^{2} \cos \chi$. The principal radii of curvature for the disk are infinite because the surface is flat. It follows that

$$
\begin{align*}
& \text { Sphere: } \sigma_{\mathrm{s}}=2 \pi a^{2}, \quad \sigma_{\text {back }}=\pi a^{2} \\
& \text { Disk: } \sigma_{\mathrm{s}}=2 \pi a^{2} \cos \chi, \quad \sigma_{\text {back }}=\infty \tag{12.5.6}
\end{align*}
$$

The infinite value might seem surprising, but its meaning becomes obvious when we recall the general interpretation of a scattering cross section as the area that blocks the incident wave. Consider the reflection from an infinite plane surface. An incident wave at normal incidence is reflected backward, so that $\bar{I}_{\mathrm{S}}=-\bar{I}_{\mathrm{I}}$ everywhere. To evaluate the scattered power in this case, we could define $\mathcal{S}_{0}$ to be a hemisphere of radius $r_{0}$ centered on the surface, then let $r_{0} \rightarrow \infty$. The result would be an infinite value of $\mathcal{P}_{s}$ because $\left|\bar{I}_{s}\right|$ is independent of $r_{0}$. Thus, the scattering cross section $\sigma$ of an infinite wall is infinite. The backscatter cross section takes the backscatter differential cross section to be representative of the scatterer's properties in all directions. In other
words, it is infinite because it does not recognize the fact that the surface has finite extent.

The preceding result may be extended to treat incidence on a surface that is locally reacting, with impedance $\rho_{0} c \zeta$. Backscatter corresponds to normal incidence, so the coefficient of reflection is

$$
\begin{equation*}
R=\frac{\zeta-1}{\zeta+1} \tag{12.5.7}
\end{equation*}
$$

The geometrical features of the reflection process are unaltered. Thus, the tube of scattered rays has the same extent as it does for the case where the surface is rigid. From the analysis in Sect.5.2.2, we know that the power flow in the reflected ray tube as it leaves the surface is $d \mathcal{P}_{s}=|R|^{2} d \mathcal{P}_{\mathrm{I}}$. The time-averaged intensity in the tube when it crosses the surrounding sphere still is $I_{\mathrm{s}}=d \mathcal{P}_{s} / d \mathcal{S}_{0}$. The result is that the previous expression for the farfield scattered intensity becomes

$$
\begin{equation*}
I_{\mathrm{s}}=I_{\mathrm{I}}|R|^{2} \frac{r_{1} r_{2}}{4 r_{0}^{2}} \tag{12.5.8}
\end{equation*}
$$

The corresponding expression for the backscatter cross section is

$$
\begin{equation*}
\sigma_{\text {back }}=\pi r_{1} r_{2}|R|^{2} \tag{12.5.9}
\end{equation*}
$$

We know that $|R|<1$ for any passive nonideal material, which means that rigid and pressure-release conditions lead to the largest backscatter at high frequencies.

The geometrical acoustics approximation may also be invoked to analyze bistatic scattering. The general approach entails formulating the KHIT with the surface response described as though the surface is locally planar. Thus, the reflected signal constitutes the scattered field on the surface. The reflection coefficient for the ray that is incident at surface point $\bar{x}_{s}$ is

$$
\begin{equation*}
R\left(\bar{x}_{s}\right)=\frac{\zeta \bar{e}_{\mathrm{I}} \cdot \bar{n}\left(\bar{x}_{s}\right)-1}{\zeta \bar{e}_{\mathrm{I}} \cdot \bar{n}\left(\bar{x}_{s}\right)+1} \tag{12.5.10}
\end{equation*}
$$

In the geometrical acoustics limit, rays are incident on the illuminated side only. Correspondingly, the pressure and the particle velocity are taken to be zero on the shadowed side. This region is identifiable by $\bar{e}_{\mathrm{I}}$ being outward from the surfaces. Thus, the total pressure and normal velocity on the surface are

$$
\begin{align*}
\bar{e}_{\mathrm{I}} \cdot \bar{n}\left(\bar{x}_{s}\right)<0 \Longrightarrow\left\{\begin{array}{l}
P\left(\bar{x}_{s}\right)=\left[1+R\left(\bar{x}_{s}\right)\right] P_{\mathrm{I}}\left(\bar{x}_{s}\right) \\
\bar{V}_{s}\left(\bar{x}_{s}\right) \cdot \bar{n}\left(\bar{x}_{s}\right)=\frac{1}{\rho_{0} c}\left[-1+R\left(\bar{x}_{s}\right)\right] P_{\mathrm{I}}\left(\bar{x}_{s}\right) \bar{e}_{\mathrm{I}} \cdot \bar{n}\left(\bar{x}_{s}\right)
\end{array}\right. \\
\bar{e}_{\mathrm{I}} \cdot \bar{n}\left(\bar{x}_{s}\right)>0 \Longrightarrow P\left(\bar{x}_{s}\right)=0 \& \bar{V}_{s}\left(\bar{x}_{s}\right) \cdot \bar{n}\left(\bar{x}_{s}\right)=0 \tag{12.5.11}
\end{align*}
$$

To implement Eq. (12.1.5), we could define $\bar{e}_{z}$ to align with $\bar{e}_{\mathrm{I}}$, and place the origin at the centroid of the body. Then, if we set the phase of the incident plane wave to zero at the origin, we would have

$$
\begin{equation*}
P_{\mathrm{I}}\left(\bar{x}_{s}\right)=\hat{P}_{\mathrm{I}} e^{-i k \bar{x}_{s} \cdot \bar{e}_{z}} \tag{12.5.12}
\end{equation*}
$$

Substitution of these representations into the KHIT would define all of the terms in the integrand, but that does not mean we can directly proceed to evaluating the integral numerically. One difficulty that awaits is the necessity to formulate a way of describing $\bar{x}_{s}$ and $\bar{n}\left(\bar{x}_{s}\right)$ in a way that fits the selected integration scheme. This is a standard issue for any formulation of KHIT. The more difficult feature stems from the fact that this is a high-frequency theory. Thus, $k \bar{x}_{s} \cdot \bar{e}_{3}$ will be much greater than unity over most of the surface. This means that the integrand of Eq. (12.1.5) would be a rapidly oscillating function of position. It is difficult to obtain a convergent result in such situations without expending a great deal of computational effort. Asymptotic methods, such as the method of stationary phase, ${ }^{4}$ might be useful for such efforts.

### 12.6 Scattering from Spheres

We have analyzed scattering from a variety of bodies in the low-frequency limit of the Rayleigh approximation and the high-frequency limit of the Kirchhoff approximation. However, we have no idea of the actual frequency range for which either approximation is valid, nor of how accurate either is in its range of validity. Answering these questions for an arbitrary scattering object would require considerable computational or experimental work. Fortunately, the spherical shape is amenable to analytical investigations. As was noted at the outset, a basic concept for an analysis of scattering is to convert the problem to an analogous radiation problem. Radiation from spheres was the subject of Sect.7.2.

### 12.6.1 Stationary Spherical Scatterer

For scattering from a sphere, there is no special orientation of the polar axis. Therefore, we align the $z$-axis with the propagation direction of the incident plane wave, $\bar{e}_{z}=\bar{e}_{\mathrm{I}}$. The origin of the coordinate system is situated at the center of the sphere, and the phase of the incident wave is defined to be zero at that location. Thus, the incident signal is $\hat{P}_{\mathrm{I}} \exp (-i k z)$. In spherical coordinates, the axial distance is $z=r \cos \psi$.

[^59]An identity known as an addition theorem ${ }^{5}$ converts the exponential to a spherical harmonic series,

$$
\begin{equation*}
P_{\mathrm{I}}=\hat{P}_{\mathrm{I}} e^{-i k z}=\hat{P}_{\mathrm{I}} \sum_{m=0}^{\infty}(2 m+1)(-i)^{m} j_{m}(k r) P_{m}(\cos \psi) \tag{12.6.1}
\end{equation*}
$$

Rigidity requires that the surface velocity be zero. The surface normal is the radial direction. Correspondingly, continuity of the normal velocity on the surface requires that

$$
\begin{equation*}
\left.\frac{\partial P_{\mathrm{s}}}{\partial r}\right|_{r=a}=-\left.\frac{\partial P_{\mathrm{I}}}{\partial r}\right|_{r=a}=-k \hat{P}_{\mathrm{I}} \sum_{m=0}^{\infty}(2 m+1)(-i)^{m} j_{m}^{\prime}(k a) P_{m}(\cos \psi) \tag{12.6.2}
\end{equation*}
$$

A spherical harmonic series for $P_{\mathrm{s}}$ satisfies the Helmholtz equation and the Sommerfeld radiation condition. The incident field is independent of the azimuthal angle $\theta$, which means that the scattered field will be axisymmetric relative to the $z$ axis. Therefore, the scattered pressure may be represented as a series of axisymmetric spherical harmonics, see Eq. (7.2.1),

$$
\begin{equation*}
P_{\mathrm{s}}=\sum_{m=0}^{\infty} B_{m} h_{m}(k r) P_{m}(\cos \psi) \tag{12.6.3}
\end{equation*}
$$

Recall that the spherical Hankel functions appearing here are the second kind, which correspond to waves that propagate outward when the convention is to use $\exp (i \omega t)$. The series coefficients are determined by using this form to satisfy the surface boundary condition. Doing so leads to

$$
\begin{equation*}
B_{m}=-(2 m+1)(-i)^{m} \frac{j_{m}^{\prime}(k a)}{h_{m}^{\prime}(k a)} \hat{P}_{\mathrm{I}} \tag{12.6.4}
\end{equation*}
$$

The corresponding expression for the scattered pressure at any location is

$$
\begin{equation*}
P_{\mathrm{s}}=-\hat{P}_{\mathrm{I}} \sum_{m=0}^{\infty}(2 m+1)(-i)^{m} \frac{j_{m}^{\prime}(k a)}{h_{m}^{\prime}(k a)} h_{m}(k r) P_{m}(\cos \psi) \tag{12.6.5}
\end{equation*}
$$

The series expansion may be specialized to the farfield by applying the asymptotic expansion in Eq. (7.1.35), which is

$$
\begin{equation*}
h_{m}(k r) \rightarrow \frac{i^{m+1}}{k r} e^{-i k r} \tag{12.6.6}
\end{equation*}
$$

[^60]This limiting form is valid if $k r \gg m$. This attribute might seem to be problematic because the spherical harmonic series sums over all $m$. Fortunately, the $B_{m}$ coefficients decrease rapidly as $m$ increases. (To verify this behavior, consider Eq. (7.1.36), which describes the behavior of the spherical functions as $m$ increases.) Thus, for any $k r$, there always is a value $m=M$ for which $B_{M}$ is sufficiently small to permit truncation of the series. We may use Eq. (12.6.6) to describe terms for $m \leq M$. We thereby obtain

$$
\begin{equation*}
P_{\mathrm{s}}=P_{\mathrm{I}} \frac{a}{r} e^{-i k r} \chi(\psi, k a), \quad \chi(\psi, k a)=-i \sum_{m=0}^{M}(2 m+1) \frac{j_{m}^{\prime}(k a)}{k a h_{m}^{\prime}(k a)} P_{m}(\cos \psi) \tag{12.6.7}
\end{equation*}
$$

One approach to evaluate the series is to accumulate the contributions of increasing $m$ until there is negligible change in the value of the sum. However, modern software is most efficient when the operations are vectorized, which requires that the range of the summation be set prior to the actual computation. A viable approach sets the maximum $m$ as an integer multiple of $k a$. If that multiple is too large, the value of $j_{m}^{\prime}(k a) / h_{m}^{\prime}(k a)$ will cause an underflow; the results reported here were obtained using $\max (m)=\operatorname{ceil}(5 k a)$. Vectorization is achieved by defining a column vector $\{F\}$ that consists of the values of $(2 m+1) j_{m}^{\prime}(k a) /\left(k a h_{m}^{\prime}(k a)\right)$ for $0 \leq m \leq \max (m)$. If the Legendre function values $P_{m}(\cos \psi)$ are stored in a column vector $\{L\}$, then the summation reduces to $\{F\}^{\mathrm{T}}\{L\}$.

The function $\chi(\psi, k a)$ is analogous to the directivity function for radiation. From it, the differential scattering cross section Eq. (12.4.9) may be evaluated as

$$
\begin{equation*}
\frac{d \sigma}{d \Gamma}=a^{2}|\chi(\psi, k a)|^{2} \tag{12.6.8}
\end{equation*}
$$

The backscatter cross section is obtained from the time-averaged intensity at $\psi=$ $180^{\circ}$, for which $P_{m}(-1)=(-1)^{m}$, so that

$$
\begin{equation*}
\chi(\pi, k a)=\sum_{m=0}^{\infty}(-1)^{m}(2 m+1) \frac{j_{m}^{\prime}(k a)}{k a h_{m}^{\prime}(k a)} \tag{12.6.9}
\end{equation*}
$$

The backscatter cross section is

$$
\begin{equation*}
\sigma_{\text {back }}=4 \pi a^{2}|\chi(\pi, k a)|^{2} \tag{12.6.10}
\end{equation*}
$$

For comparison, the value of $\sigma_{\text {back }}$ in the Rayleigh scattering limit may be found from Eq. (12.3.22) by setting $W_{3,3}=2 \pi a^{3}, \psi_{\mathrm{I}}=0$, and $\psi_{0}=-\pi$, whereas the geometrical acoustics limit is independent of $k a$,

$$
\begin{gather*}
\sigma_{\text {back }}=\frac{25}{9} \pi a^{2}(k a)^{4}, \quad k a \ll 1  \tag{12.6.11}\\
\sigma_{\text {back }}=2 \pi a^{2}, \quad k a \gg 1
\end{gather*}
$$

In Fig. 12.7, the computed backscatter cross section is compared to the low- and high-frequency approximations. The lower graph shows that the Rayleigh limit is indistinguishable from the spherical harmonic series up to $k a=0.6$, while the geometrical acoustic approximation essentially is close to the mean about which the computed value fluctuates when $k a>10$. The amplitude of this fluctuation decreases as $k a$ increases, with an excursion that is less than $5 \%$ of the mean value above $k a=20$.


Fig. 12.7 Frequency dependence of the backscatter cross section of a rigid, stationary sphere

Evaluation of the scattering cross section requires integration of the time-averaged bistatic intensity. This quantity in the farfield is $\bar{e}_{r}\left|P_{\mathrm{s}}\right|^{2} /\left(\rho_{0} c\right)$. It contains all cross products of spherical harmonics. However, an evaluation of radiated power entails integration over the surrounding sphere, and the Legendre functions are orthogonal over that domain, see Eq. (7.1.16). Orthogonality was shown in Eq. (7.1.17) to decouple the spherical harmonics in regard to evaluation of radiated power. In the present context, the radiated power in the scattered field is

$$
\begin{align*}
\mathcal{P}_{\mathrm{s}} & =\int_{0}^{\pi} \frac{P_{\mathrm{s}} P_{\mathrm{s}}^{*}}{2 \rho_{0} c}\left(2 \pi r^{2} \sin \psi d \psi\right) \\
& =\frac{\left|P_{\mathrm{I}}\right|^{2}}{2 \rho_{0} c} \sum_{m=0}^{\infty} 4 \pi a^{2}(2 m+1)\left(\frac{j_{m}^{\prime}(k a)}{k a\left|h_{m}^{\prime}(k a)\right|}\right)^{2} \tag{12.6.12}
\end{align*}
$$

The factor preceding the summation is the intensity of the incident wave, so the scattering cross section of a rigid sphere is

$$
\begin{equation*}
\sigma_{\mathrm{s}}=\frac{\mathcal{P}_{\mathrm{s}}}{I_{\mathrm{I}}}=4 \pi a^{2} \sum_{m=0}^{\infty} 2(2 m+1)\left(\frac{j_{m}^{\prime}(k a)}{k a\left|h_{m}^{\prime}(k a)\right|}\right)^{2} \tag{12.6.13}
\end{equation*}
$$

Figure 12.8 shows that unlike the backscatter cross section, $\sigma_{\mathrm{s}}$ monotonically increases toward the geometrical acoustics limit, $\sigma_{\mathrm{s}}=2$, without any fluctuation. The value of $\sigma_{\mathrm{s}}$ in the Rayleigh limit, see Eq. (12.4.6), closely matches the computed function for $k a<0.45$.


Fig. 12.8 Frequency dependence of the scattering cross section of a rigid, stationary sphere

Before we examine why $\sigma_{\text {back }}$ and $\sigma_{\mathrm{s}}$ behave differently, let us examine the scattering response from the perspective of the scattering directivity function $\chi(\psi, k a)$. Graphs like those in Fig. 12.9 were published by Stenzel in 1938, ${ }^{6}$ which is quite impressive given the nonexistence of electronic computing. At low frequencies, the angular dependence of $\chi$ fits the Rayleigh limit in Eq. (12.3.22), with the monopole and dipole contributions being in-phase for $\psi>90^{\circ}$ and oppositely phased for $\psi<90^{\circ}$. As the frequency increases, side lobes become increasingly numerous,


Fig. 12.9 Scattering directivity function $\chi(\psi, k a)$ as a function of the observation angle relative to the propagation direction $\bar{e}_{\mathrm{I}}$ of the incident wave

[^61]which is a trend for acoustic radiation problems. An unexpected trend that grows as the frequency increases is strong forward scattering, in which the maximum value of $\chi$ increases and the beamwidth decreases as $k a$ increases. This behavior is not predicted by ray theory, according to which the incident signal should be blocked in the shadow of the sphere. Because $P=P_{\mathrm{s}}+P_{I}$, such reasoning leads to the expectation that forward scattering should be such that $\left|P_{\mathrm{s}}\right|=\left|P_{\mathrm{I}}\right|$.

The inability of geometrical acoustics to predict forward scattering is not a failure of the theory. This phenomenon is another diffraction effect. Such effects arose in our study of geometrical acoustics, which does not predict that sound will be heard on the quiet side of a caustic. It also arose there as the fact that a ray at critical incidence to the interface of two fluid might propagate along the interface, then re-emerge into the transmitting fluid. The term "diffraction" refers to any high-frequency phenomenon that cannot be described by ray acoustics. The process that ray acoustics missed in scattering is creeping waves. Their existence would be difficult to detect from direct evaluations of the spherical harmonic description of the pressure field. At the same time, availability of a series solution is crucial, because it is the beginning point for a mathematical proof of the existence of creeping rays.

Watson in $1918^{7}$ addressed the fact that series such as those derived here require an increased number of terms to converge as the frequency increases. Only rudimentary computational tools were available, so he needed to be "smart". He developed a technique that transforms the series to a contour integral. Evaluation by the methods of complex analysis produces a solution in the form of a series of residues. Each residue corresponds to a pair of waves that pass a point on the surface. Figure 12.10 provides a qualitative picture of the first pair, which are the dominant contributors to the signal.

The incident rays depicted in Fig. 12.10 touch the sphere at $\psi=90^{\circ}$, where they are tangent to the sphere. In other words, they are at grazing incidence. They launch rays that propagate along meridian circles (constant $\theta$ ) in the sense of decreasing $\psi$, which is the reason they are said to creep over the surface. Ultimately, they focus at the shadowed pole $(\psi=0)$. They emerge from there as waves that travel in the sense of increasing $\psi$ along the diametrically opposite meridian, for which the azimuthal

Fig. 12.10 Incident rays $P_{\mathrm{I}}$ that are at grazing incidence curve around the sphere as creeping waves (the dotted lines). They circumnavigate the sphere along meridians


[^62]angle is $-\theta$. In other words, an individual ray of the creeping wave follows a great circle that contains the polar axis. Each ray circumnavigates the sphere along its great circle. Eventually, they return to the location where they were launched. From there, the rays execute another circumnavigation, then another, and so on. Each passage is described in Watson's solution by another residue in the complex integration.

The reason it is sufficient to only account for the first circumnavigation is that the wavenumber for these rays is complex, so they attenuate as they propagate. Because this is a high-frequency phenomenon, the behavior in the vicinity of any surface patch is like what it would be if that patch laid in a planar boundary. Thus, the surface ray launches a tangential ray. This signal is said to be shed from the surface. The shedding at a few surface locations is depicted in Fig. 12.11, but it occurs continuously along the surface. These shed waves are the agents that illuminate the shadow zone.

Fig. 12.11 Rays shed from creeping wave on the surface of a sphere


The wave that is shed in the backscattering direction combines with the reflected signal predicted by geometrical acoustics. The latter is referred to as the specularly reflected wave. The fluctuations of $\sigma_{\text {back }}$ displayed in Fig. 12.7 stem from constructive and interference between the contribution of the creeping waves and the specularly reflected wave. As the frequency increases, the decay rate for the creeping waves increases. Above $k a=20$, this attenuation is sufficiently strong to make shedding in the backscatter direction negligible. This leaves only the contribution of specular reflection, which explains why the computed $\sigma_{\text {back }}$ differs little from the geometrical acoustic limit at very high frequencies.

Diffraction is responsible for acoustical phenomena in other systems. Sometimes, we can identify the mechanism directly from a mathematical analysis. For example, the side lobes of a sound beam are a diffraction effect associated with the discontinuity of the velocity at the edge of a piston in a baffle, which leads to center and edge waves for the on-axis signal. In other situations, specialized analyses based on $k a$ being very large are required. This is the situation regarding the Watson transformation. The mathematical sophistication required to carry out such analyses exceeds the level set for this text. An excellent introduction is available in the text by Junger and Feit. ${ }^{8}$ Pierce ${ }^{9}$ explores the same phenomena, as well as diffraction around wedges

[^63]and in ocean acoustics, and Keller's geometrical theory of diffraction ${ }^{10}$ offers much in terms of comprehensiveness and generality.

### 12.6.2 Scattering by an Elastic Spherical Shell

Acoustic radiation from a spherical shell was shown in Sect. 7.2.4 to be profoundly dependent on frequency. At the resonance frequencies, the sole agent sustaining any excitation applied to the shell is the resistive part of the acoustic loading. Now, consider the implications for scattering from a spherical shell. The surface pressure in the incident signal may be regarded conceptually as the excitation, so the pressure in the scattered signal is analogous to the radiated field. Thus, if the frequency is close to one of these fluid-loaded resonances, it is reasonable to expect that the field will be quite different from what it would be away from these frequencies. It also is reasonable to expect that if the shell is composed of a material whose density and sound speed are much greater than the fluid's properties, the scattered signal away from resonances would be like that for a rigid target. Our objective is to explore the correctness of these expectations.

We will analyze scattering from an elastic spherical shell by merging the analyses of scattering from a rigid sphere and radiation from a spherical shell. Toward that end, let us decompose the scattered field in two parts. The first is that which would be found if the shell was rigid and stationary. It is called the blocked pressure, denoted as $p_{\mathrm{b}}=\operatorname{Re}\left(P_{\mathrm{b}} \exp (i \omega t)\right)$. The second contribution is induced by the motion of the shell resulting from its elasticity. Correspondingly, we will denote this part of the scattered field as $p_{\mathrm{e}}=\operatorname{Re}\left(P_{\mathrm{e}} \exp (i \omega t)\right)$. Thus, we have

$$
\begin{equation*}
P_{\mathrm{s}}=P_{\mathrm{b}}+P_{\mathrm{e}} \tag{12.6.14}
\end{equation*}
$$

This decomposition is substituted into Eq. (12.1.3), which govern the scattered field. When each term is associated with its respective effect, we obtain

$$
\begin{align*}
\nabla^{2} P_{\mathrm{b}}+k^{2} P_{\mathrm{b}} & =0, \quad \frac{\partial P_{\mathrm{b}}}{\partial r}=-\frac{\partial P_{\mathrm{I}}}{\partial r} \text { at } r=a  \tag{12.6.15a}\\
\nabla^{2} P_{\mathrm{e}}+k^{2} P_{\mathrm{e}} & =0, \quad \frac{\partial P_{\mathrm{e}}}{\partial r}=\omega^{2} \rho_{0} W \text { at } r=a \tag{12.6.15b}
\end{align*}
$$

In the boundary condition for $P_{\mathrm{e}}$, the parameter $W$ is the complex amplitude of the radial displacement of the shell, that is, $w=\operatorname{Re}(W \exp (i \omega t))$. The corresponding radial velocity is $v_{r}=\operatorname{Re}(i \omega W \exp (i \omega t))$. In addition to satisfying the Helmholtz equation and a surface boundary condition, both $P_{\mathrm{b}}$ and $P_{\mathrm{e}}$ must satisfy the Sommerfeld radiation condition, Eq. (12.1.4).

[^64]The conditions that the blocked pressure must satisfy are identical to those for the pressure scattered by a rigid sphere. Our interest is the case of an incident plane wave, which was analyzed in the previous section. Consequently, we may employ Eqs. (12.6.3) and (12.6.4) to describe the blocked pressure, so that

$$
\begin{equation*}
P_{\mathrm{b}}=-\hat{P}_{\mathrm{I}} \sum_{m=0}^{\infty}(2 m+1)(-i)^{m} \frac{j_{m}^{\prime}(k a)}{h_{m}^{\prime}(k a)} h_{m}(k r) P_{m}(\cos \psi) \tag{12.6.16}
\end{equation*}
$$

The corresponding farfield description is provided by Eq. (12.6.7).
Now, let us analyze $P_{\mathrm{e}}$. The radial displacement $w$ represents the response of the elastic shell to the total acoustic pressure applied at its surface. This is the fully coupled problem solved in Sect. 7.2.4. That analysis was limited to excitations that are axisymmetric, which is true for an incident plane wave. It was shown that the spherical harmonic expansions of the radial displacement $w$ and meridional displacement $u$ in Eq. (7.2.35) will satisfy the equations of motion if the coefficients are selected properly. The specific expressions are

$$
\begin{gather*}
w=\operatorname{Re}\left(W e^{i \omega t}\right), \quad W=\sum_{m=0}^{\infty} W_{m} P_{m}(\cos \psi) \\
u=\operatorname{Re}\left(U e^{i \omega t}\right), \quad U=\sum_{m=0}^{\infty} U_{m} \frac{d}{d \psi} P_{m}(\cos \psi) \tag{12.6.17}
\end{gather*}
$$

Furthermore, the pressure induced by the shell's motion constitutes a radiated field, so it too must be expressible as a spherical harmonic series,

$$
\begin{equation*}
P_{\mathrm{e}}=\sum_{m=0}^{\infty} C_{m} P_{m}(\cos \psi) h_{m}(k r) \tag{12.6.18}
\end{equation*}
$$

Satisfaction of the velocity boundary condition, Eq. (12.6.15b), requires that

$$
\begin{equation*}
k C_{m} h_{m}^{\prime}(k a)=\omega^{2} \rho_{0} W_{m} \tag{12.6.19}
\end{equation*}
$$

Thus, $P_{\mathrm{e}}$ will be determined when we find an expression for $W_{m}$.
The total pressure acting on the surface of the shell is the sum of the incident, blocked, and structural contributions. In the basic differential equations of motion for a spherical shell, the surface pressure is decomposed into two parts: an applied loading that does not depend on the shell's motion, and an acoustic part that consists of the pressure induced by the motion. In the present context, the part that is generated by the shell's motion is $P_{\mathrm{e}}$, and the effective loading is the sum of $P_{\mathrm{I}}$ and $P_{\mathrm{b}}$. The series for $P_{\mathrm{e}}$ is the same as the description of the acoustic pressure obtained from Eq. (7.2.41), so we may determine $P_{\mathrm{e}}$ directly from the radiation analysis. To do so, we equate the applied pressure $q_{\text {applied }}$ to the negative of the sum of the incident
and blocked pressures, where the negative sign applies because $q_{\text {applied }}$ was taken to be positive if it is outward. Equation (12.6.1) is a spherical harmonic series for a plane wave propagating in the direction of increasing $z$. When it is combined with Eq. (12.6.16) for $P_{\mathrm{b}}$, the total pressure loading is

$$
\begin{align*}
q_{\text {applied }} & =-\left.\left(P_{\mathrm{I}}+P_{\mathrm{b}}\right)\right|_{r=a} \\
& =-\hat{P}_{\mathrm{I}} \sum_{m=0}^{\infty}(2 m+1)(-i)^{m}\left[j_{m}(k a)-\frac{j_{m}^{\prime}(k a)}{h_{m}^{\prime}(k a)} h_{m}(k a)\right] P_{m}(\cos \psi) \tag{12.6.20}
\end{align*}
$$

When the fraction in the bracketed term is cleared, the numerator is a quantity called a Wronskian. Application of the definition that $h_{m}(k a)=j_{m}(k a)-i n_{m}(k a)$ gives a form that is described by a simplifying identity, ${ }^{11}$

$$
\begin{align*}
j_{m}(k a) h_{m}^{\prime}(k a)-j_{m}^{\prime}(k a) h_{m}(k a) \equiv & -i\left[j_{m}(k a) n_{m}^{\prime}(k a)\right. \\
& \left.-j_{m}^{\prime}(k a) n_{m}(k a)\right]=\frac{-i}{(k a)^{2}} \tag{12.6.21}
\end{align*}
$$

The result is that the applied load reduces to

$$
\begin{equation*}
q_{\text {applied }}=-\hat{P}_{\mathrm{I}} \sum_{m=0}^{\infty} \frac{(-i)^{m+1}(2 m+1)}{(k a)^{2} h_{m}^{\prime}(k a)} P_{m}(\cos \psi) \tag{12.6.22}
\end{equation*}
$$

The radiation analysis derived algebraic expressions for the displacement coefficients $W_{m}$ and $U_{m}$ in terms of the coefficient $F_{m}$ of a series representation of the applied pressure. The preceding equation is such a description, so the force coefficients are

$$
\begin{equation*}
F_{m}=-\hat{P}_{\mathrm{I}} \frac{(-i)^{m+1}(2 m+1)}{(k a)^{2} h_{m}^{\prime}(k a)} \tag{12.6.23}
\end{equation*}
$$

The relation between $F_{m}$ and $W_{m}$ is described by Eq. (7.2.50) in terms of a structural stiffness $\left(Z_{\mathrm{e}}\right)_{m}$ and an acoustic impedance $\left(Z_{\mathrm{f}}\right)_{m}$, such that

$$
\begin{equation*}
\left[\left(Z_{\mathrm{e}}\right)_{m}+\left(Z_{\mathrm{f}}\right)_{m}\right]\left(i k_{\mathrm{e}} a\right) \frac{W_{m}}{a}=\frac{a}{h} \frac{F_{m}}{\rho_{e} c_{e}^{2}} \tag{12.6.24}
\end{equation*}
$$

where $\rho_{\mathrm{e}}$ and $c_{\mathrm{e}}$ are the density and a certain wave speed for the shell's material, and $k_{\mathrm{e}}=\omega / c_{\mathrm{e}}$. The structural impedance describes the combination of inertial and elastic effects,

[^65]\[

$$
\begin{equation*}
\left(Z_{\mathrm{e}}\right)_{m}=\frac{k_{\mathrm{e}} a}{i} \frac{\left[\left(\kappa_{m}^{(1)}\right)^{2} /\left(k_{\mathrm{e}} a\right)^{2}-1\right]\left[\left(\kappa_{m}^{(2)}\right)^{2} /\left(k_{\mathrm{e}} a\right)^{2}-1\right]}{\left(\kappa_{m}^{(0)}\right)^{2} /\left(k_{\mathrm{e}} a\right)^{2}-1} \tag{12.6.25}
\end{equation*}
$$

\]

In this expression $\kappa_{m}^{(j)}=\Omega_{m}^{(j)} a / c_{\mathrm{e}}$, where $\Omega_{m}^{(1)}$ and $\Omega_{m}^{(2)}$ are the natural frequencies for axisymmetric-free vibration of the shell in the $m$ th spherical harmonic. Also, $\kappa_{m}^{(0)}$ is a nondimensionalization of the frequency $\Omega_{m}^{(0)}$ at which there is no radial displacement, regardless of the excitation. The determination of these parameters is described in Sect.7.2.4. The parameter $k_{\mathrm{e}}=\omega / c_{\mathrm{e}}$, where $c_{\mathrm{e}}$ is the phase speed of an extensional wave of plane strain. The value $\left(Z_{\mathrm{e}}\right)_{m}$ is purely imaginary, so it is a reactance. Inclusion of the effects of internal dissipation would add a resistive part. However, this resistance would be much less than the reactance, except in the vicinity of either natural frequency, where the reactive part is zero.

The fluid impedance is

$$
\begin{equation*}
\left(Z_{\mathrm{f}}\right)_{m}=-i \frac{\rho_{0} c}{\rho_{e} c_{e}}\left(\frac{a}{h}\right) \frac{h_{m}(k a)}{h_{m}^{\prime}(k a)} \tag{12.6.26}
\end{equation*}
$$

The real part is positive, representing resistance associated with radiation damping. The imaginary part also is positive, and therefore an inertance, corresponding to an added mass effect. Figure 7.7 indicates that the reactive part has a broad peak that occurs at a frequency that increases monotonically with increasing $m$. In contrast, the resistive part of $\left(Z_{\mathrm{f}}\right)_{m}$ approaches a constant value with increasing frequency.

Equation (7.2.41) describes the coefficients of a spherical harmonic series for the pressure. Substitution of the expressions for $W_{m}$ and $F_{m}$, followed by synthesis of the series, leads to

$$
P_{\mathrm{e}}=\hat{P}_{\mathrm{I}} \frac{\rho_{0} c}{\rho_{\mathrm{e}} c_{\mathrm{e}}} \frac{a}{h} \sum_{m=0}^{\infty} \frac{(-i)^{m}(2 m+1)}{\left[\left(Z_{\mathrm{e}}\right)_{m}+\left(Z_{\mathrm{f}}\right)_{m}\right]} \frac{h_{m}(k r)}{(k a)^{2} h_{m}^{\prime}(k a)^{2}} P_{m}(\cos \psi)
$$

The farfield version is obtained by applying Eq. (12.6.6), which is the asymptotic property of the Hankel function. The result is

$$
\begin{equation*}
\left(P_{\mathrm{e}}\right)_{\mathrm{ff}}=i \hat{P}_{\mathrm{I}} \frac{\rho_{0} c}{\rho_{\mathrm{e}} c_{\mathrm{e}}} \frac{a}{h} \frac{a}{r} e^{-i k r} \sum_{m=0}^{M} \frac{(2 m+1)}{\left[\left(Z_{\mathrm{e}}\right)_{m}+\left(Z_{\mathrm{f}}\right)_{m}\right](k a)^{3} h_{m}^{\prime}(k a)^{2}} P_{m}(\cos \psi) \tag{12.6.27}
\end{equation*}
$$

where $M$ is the largest index for which the factor of the Legendre function is significant.

The amplitude of the $m$ th spherical harmonic is maximized when the magnitude of $\left(Z_{\mathrm{e}}\right)_{m}+\left(Z_{\mathrm{f}}\right)_{m}$ is a minimum. Here, an analogy with standard vibration theory comes into play. The portion $\left(Z_{\mathrm{e}}\right)_{m}$ represents the effect of inertia and elasticity. If internal dissipation is ignored, $\left(Z_{\mathrm{e}}\right)_{m}$ is imaginary. The fluid effect is contained in $\left(Z_{\mathrm{f}}\right)_{m}$. It is complex, with the imaginary part acting as a virtual mass that adds to $\left(Z_{\mathrm{e}}\right)_{\mathrm{m}}$. The real part represents radiation damping, in which power is radiated into the fluid. In a standard spring-mass-dashpot system, a resonance is associated with the magnitude of the combined impedance being a minimum. However, if the damping is not too large, the resonance condition will be close to that for which the imaginary part is zero. Similarly, although the true maximum of a spherical harmonic corresponds to the frequency that minimizes $\left|\left(Z_{\mathrm{e}}\right)_{m}+\left(Z_{\mathrm{f}}\right)_{m}\right|$, it is simpler to approximate the frequency at which a fluid-loaded resonance occurs as the value for which $\operatorname{Im}\left(\left(Z_{\mathrm{e}}\right)_{m}+\left(Z_{\mathrm{f}}\right)_{m}\right)=0$.

These resonances affect the selection of the harmonic $M$ at which the series may be truncated. In the analysis of scattering from a rigid sphere, the criterion for selection of $M$ was found to be $M \gg k a$, whose satisfaction assures that $h_{m}^{\prime}(k a)^{2}$ is very large. That criterion applies here also, but there is another consideration. It is imperative that any harmonic $\tilde{m}$ that is close to resonance be included in the summation. That is, we should set $M>\tilde{m}$. Figure 7.7 describes the impedances as a function of frequency for fixed $m$, but the issue here is the dependence of $\operatorname{Im}\left[\left(Z_{\mathrm{e}}\right)_{m}+\left(Z_{\mathrm{f}}\right)_{m}\right]$ on $m$ at fixed $k a$.

The natural frequencies as functions of $m$ are described in Fig. 7.5, and $\Omega_{m}^{(0)}$ falls between these frequencies. If $M$ is set well above the intersection of the line at the current frequency $k a$ with the lower branch, then $\operatorname{Im}\left(Z_{\mathrm{e}}\right)_{m}$ will be large negatively. This accelerates the convergence rate of the series for $P_{\mathrm{e}}$, beyond what is attributable to the behavior of the Hankel function. The next example uses $M=\operatorname{ceil}(10 \mathrm{ka})$, which is much larger than necessary.

Elastic effects occur in a number of ways in Eq. (12.6.27), so let us evaluate $P_{\mathrm{e}}$ in the case where $\operatorname{Im}\left[\left(Z_{\mathrm{e}}\right)_{\tilde{m}}+\left(Z_{\mathrm{f}}\right)_{\tilde{m}}\right]$ actually is zero. Doing so will lead to the pressure at a fluid-loaded resonance. The frequency for this resonance is denoted as $\tilde{\omega} \equiv \tilde{k} c$. The total impedance reduces to $\operatorname{Re}\left(Z_{\mathrm{f}}\right)_{\tilde{m}}$ at $\tilde{k} a$. Equation (12.6.26) describes the fluid impedance, so we have

$$
\begin{equation*}
\left.\left[\left(Z_{\mathrm{e}}\right)_{\tilde{m}}+\left(Z_{\mathrm{f}}\right)_{\tilde{m}}\right]\right|_{\tilde{k} a}=\left.\operatorname{Re}\left(Z_{\mathrm{f}}\right)_{\tilde{m}}\right|_{\tilde{k} a}=\frac{\rho_{0} c a}{\rho_{\mathrm{e}} c_{\mathrm{e}} h} \operatorname{Re}\left[-i \frac{h_{\tilde{m}}(\tilde{k} a)}{h_{\tilde{m}}^{\prime}(\tilde{k} a)}\right] \tag{12.6.28}
\end{equation*}
$$

The Hankel function of the second kind is defined in terms of the real spherical Bessel and Neumann functions to be $h_{\tilde{m}}^{(2)}(\tilde{k} a)=j_{\tilde{m}}(\tilde{k} a)-i n_{\tilde{m}}(\tilde{k} a)$. With this, the impedance becomes

$$
\begin{align*}
{\left.\left[\left(Z_{\mathrm{e}}\right)_{\tilde{m}}+\left(Z_{\mathrm{f}}\right)_{\tilde{m}}\right]\right|_{\tilde{k} a} } & =\frac{\rho_{0} c a}{\rho_{\mathrm{e}} c_{\mathrm{e}} h} \operatorname{Re}\left[-i \frac{h_{\tilde{m}}(\tilde{k} a) h_{\tilde{m}}^{\prime}(\tilde{k} a)^{*}}{h_{\tilde{m}}^{\prime}(\tilde{k} a) h_{\tilde{m}}^{\prime}(\tilde{k} a)^{*}}\right] \\
& =\frac{\rho_{0} c a}{\rho_{\mathrm{e}} c_{\mathrm{e}} h}\left(\frac{-n_{\tilde{m}}(\tilde{k} a) j_{\tilde{m}}^{\prime}(\tilde{k} a)+j_{\tilde{m}}(\tilde{k} a) y_{\tilde{m}}^{\prime}(\tilde{k} a)}{\left|h_{\tilde{m}}^{\prime}(\tilde{k} a)\right|^{2}}\right) \tag{12.6.29}
\end{align*}
$$

The numerator is the Wronskian in Eq. (12.6.21), so the total impedance at a fluidloaded resonance reduces to

$$
\begin{equation*}
\left.\left[\left(Z_{\mathrm{e}}\right)_{m}+\left(Z_{\mathrm{f}}\right)_{\tilde{m}}\right]\right|_{\tilde{k} a}=\frac{\rho_{0} c a}{\rho_{\mathrm{e}} c_{\mathrm{e}} h} \frac{1}{(\tilde{k} a)^{2}\left|h_{m}^{\prime}(\tilde{k} a)\right|^{2}} \tag{12.6.30}
\end{equation*}
$$

Presumably, the $\tilde{m}$ term in $\left(P_{\mathrm{e}}\right)_{\mathrm{ff}}$ is dominant. Substitution of the impedance at resonance into Eq. (12.6.27) yields

$$
\begin{equation*}
\left(P_{\mathrm{e}}\right)_{\mathrm{ff}} \approx \hat{P}_{\mathrm{I}} \frac{e^{-i \tilde{k} r}}{\tilde{k} r}(2 \tilde{m}+1) e^{-2 i \arg \left(h_{\tilde{m}}^{\prime}(\tilde{k} a)\right)} P_{\tilde{m}}(\cos \psi) \tag{12.6.31}
\end{equation*}
$$

The surprising aspect of this result is that the peak magnitude of $\left(P_{\mathrm{e}}\right)_{\mathrm{ff}} / P_{\mathrm{I}}$ appears to depend only on the harmonic number and the acoustic wavenumber at which the resonance occurs. However, this is a deceptive view, because both $\tilde{m}$ and $\tilde{k}$ depend strongly on all system parameters.

EXAMPLE 12.4 A spherical shell $(h / a=40)$ composed of aluminum ( $\rho_{\mathrm{e}}=$ $2100 \mathrm{~kg} / \mathrm{m}^{3}, c_{\mathrm{e}}=6410 \mathrm{~m} / \mathrm{s}$ ) is submerged in water. Determine the backscatter cross section $\sigma_{\text {back }}$ and total scattering cross section $\sigma$ in the range $k a<10$. Compare this to the frequency dependence of each parameter if the sphere was rigid and stationary. Then, repeat the analysis for the case where the sphere is suspended in the air.

## Significance

The dependence on a number of system parameters, as well as the interplay between spherical harmonic number and frequency, can best be understood by performing quantitative analyses like the one to be performed here. The evaluation itself will provide a useful review of some fundamental concepts.

## Solution

The total scattered field is the sum of the blocked and structural contributions. Because $P_{\mathrm{b}}$ is the same as the scattered field for the rigid case, we shall compute it, and
$P_{\mathrm{e}}$ as separate entities. Our strategy is rather straightforward. The in-vacuo natural frequencies $\Omega_{m}^{(1)}$ and $\Omega_{m}^{(2)}$, as well as the zero-radial-displacement frequency $\Omega_{m}^{(0)}$, are independent of the excitation frequency. The truncation criterion is $M=10 \mathrm{ka}$, and the maximum frequency to be considered is $k a=10$. Therefore, we begin by computing the modal properties for azimuthal harmonics from zero to 100 . This computation entails solving the characteristic equation, Eq. (7.2.44). The $K$ coefficients to use are those in Eq. (7.2.48), which describe membrane and flexural deformation effects. The roots of this quadratic equation are the values $\kappa_{m}^{(j)}=\Omega_{m}^{(j)} a / c_{\mathrm{e}}, j=1,2$, and $\kappa_{m}^{(0)}$ in Eq. (7.2.46).

The remaining operations are performed for a succession of $k a$ value. The values of $h_{m}(k a)$ for $0 \leq m \leq M+1$ are computed and used in conjunction with a recurrence relation to evaluate $h_{m}^{\prime}(k a)$ for $0 \leq m \leq M$. These values are required to evaluate both $P_{\mathrm{b}}$ and $P_{\mathrm{e}}$, and the latter also requires $j_{m}^{\prime}(k a)$, which is the real part of $h_{m}^{\prime}(k a)$. Let us denote the coefficients of the spherical harmonic series for $P_{\mathrm{b}} / P_{\mathrm{I}}$ in the farfield as $C_{\mathrm{b}, m}$ and those for $P_{\mathrm{e}} / P_{\mathrm{I}}$ as $C_{\mathrm{e}, m}$. Thus, from Eqs. (12.6.7) and (12.6.27), we have

$$
\begin{gather*}
\frac{P_{\mathrm{b}}}{P_{\mathrm{I}}}=\frac{a}{r} e^{-i k r} \sum_{m=0}^{M} C_{\mathrm{b}, m} P_{m}(\cos \psi), \quad \frac{P_{\mathrm{e}}}{P_{\mathrm{I}}}=\frac{a}{r} e^{-i k r} \sum_{m=0}^{M} C_{\mathrm{e}, m} P_{m}(\cos \psi) \\
C_{\mathrm{b}, m}=-i(2 m+1) \frac{j_{m}^{\prime}(k a)}{k a h_{m}^{\prime}(k a)}  \tag{1}\\
C_{\mathrm{e}, m}=i \frac{\rho_{0} c a}{\rho_{\mathrm{e}} c_{\mathrm{e}} h} \frac{(2 m+1)}{\left[\left(Z_{\mathrm{e}}\right)_{m}+\left(Z_{\mathrm{f}}\right)_{m}\right](k a)^{3} h_{m}^{\prime}(k a)^{2}}
\end{gather*}
$$

All coefficients for a specific $k a$ may be computed in a single vectorized step.
The backscatter cross section was defined in Eq. (12.4.10). The backscatter direction is $\psi=\pi$, for which $P_{m}(-1)=(-1)^{m}$. Thus,

$$
\begin{gather*}
\left(\sigma_{\text {back }}\right)_{\mathrm{e}}=4 \pi a^{2}\left|\sum_{m=0}^{M}(-1)^{m}\left(C_{\mathrm{b}, m}+C_{\mathrm{e}, m}\right)\right|^{2} \\
\left(\sigma_{\text {back }}\right)_{\mathrm{b}}=4 \pi a^{2}\left|\sum_{m=0}^{M}(-1)^{m} C_{\mathrm{b}, m}\right|^{2} \tag{2}
\end{gather*}
$$

Evaluation of the scattering cross section would be more difficult were it not for the orthogonality property of Legendre functions. Because of it, there is no coupling of harmonics in the scattered power. The steps for this analysis are

$$
\begin{align*}
\sigma_{\mathrm{s}} & =\frac{\mathcal{P}_{\mathrm{s}}}{I_{\mathrm{I}}}=\int_{0}^{\pi}\left(\frac{P_{\mathrm{b}}}{P_{\mathrm{I}}}+\frac{P_{\mathrm{w}}}{P_{\mathrm{I}}}\right)\left(\frac{P_{\mathrm{b}}^{*}}{P_{\mathrm{I}}^{*}}+\frac{P_{\mathrm{w}}^{*}}{P_{\mathrm{I}}^{*}}\right)\left(2 \pi r^{2} \sin \psi d \psi\right) \\
& =2 \pi a^{2} \sum_{j=0}^{M} \sum_{m=0}^{M} \int_{0}^{\pi}\left(C_{\mathrm{b}, j}+C_{\mathrm{e}, j}\right)\left(C_{\mathrm{b}, m}^{*}+C_{\mathrm{e}, m}^{*}\right) P_{j}(\cos \psi) P_{m}(\cos \psi) \sin \psi d \psi \\
& =\pi a^{2} \sum_{m=0}^{M} \frac{4}{2 m+1}\left[\left|C_{\mathrm{b}, m}\right|^{2}+\left|C_{\mathrm{e}, m}^{*}\right|^{2}+2 \operatorname{Re}\left(C_{\mathrm{b}, m} C_{\mathrm{e}, m}^{*}\right)\right] \tag{3}
\end{align*}
$$

The scattering cross section for the rigid sphere is obtained by setting $C_{\mathrm{e}, m}=0$.
Figure 1 shows the backscatter and total scattering cross sections. The spikes correspond to fluid-loaded resonances, whose occurrence was anticipated in the earlier discussion. The puzzling aspect is their occurrence only in the low-frequency band.


Figure 1

An examination of this behavior begins by considering the resonance condition in the absence of fluid loading. In that case, $\left(Z_{f}\right)_{\tilde{m}}$ is identically zero, so a resonance corresponds to $\left(Z_{\mathrm{e}}\right)_{m}$, which is imaginary, being zero. At a randomly selected $k a$, a plot of $\operatorname{Im}\left(Z_{\mathrm{e}}\right)_{m}$ as a function of $m$ would show that it crosses the zero axis at a value of $m$ that is not an integer, so it is not a resonant frequency. Increasing $k a$ eventually will bring on a condition where $\operatorname{Im}\left(Z_{e}\right)_{m}=0$ at an integer value of $m$. Because we are ignoring the role of fluid loading, the corresponding value of $\left(C_{\mathrm{e}}\right)_{\tilde{m}}$ would be infinite. Further increase of $k a$ increases the $m$ value at which $\left(Z_{\mathrm{e}}\right)_{m}=0$, until it reaches another integer value of $m$, and so on.

A crucial aspect of the effect of fluid loading is that $\left(Z_{\mathrm{f}}\right)_{m}$ is zero only at $k a=0$. The two quantities that affect the magnitude of $\left(C_{\mathrm{e}}\right)_{m}$ are $\left|\left(Z_{\mathrm{e}}\right)_{m}+\left(Z_{\mathrm{f}}\right)_{m}\right|$ and $\left|h_{m}^{\prime}(k a)\right|$, both of which are in the denominator of the above expression. A fluidloaded resonance was defined as a frequency $\tilde{k} a$ at which $\operatorname{Im}\left(\left(Z_{\mathrm{e}}\right)_{\tilde{m}}+\left(Z_{\mathrm{f}}\right)_{\tilde{m}}\right)=0$. However, $\operatorname{Re}\left(Z_{\mathrm{f}}\right)_{\tilde{m}}$ is not zero, so it is possible that $\left|\left(Z_{\mathrm{e}}\right)_{m}+\left(Z_{\mathrm{f}}\right)_{m}\right|$ is not close to a minimum. Furthermore, even if it is, the denominator of $\left(C_{\mathrm{e}}\right)_{\tilde{m}}$ might not be
minimized because the magnitude of its denominator is $\left|\left(Z_{\mathrm{e}}\right)_{m}+\left(Z_{\mathrm{f}}\right)_{m}\right|\left|h_{m}^{\prime}(k a)\right|$; a small impedance might be multiplied by a large Hankel function.

To see how these factors work together, Fig. 2 examines them at $k a=1.188$, at which the first peak in $\sigma_{\text {back }}$ and $\sigma_{\mathrm{s}}$ occur. This value of $k a$ is very close to the first natural frequency of the $m=2$ spherical harmonic, $k a \approx \kappa_{2}^{(1)}$. The uppermost graph indicates that the reactance at $m=2$ is zero, and the resistance is small. The actual values are $\left(Z_{\mathrm{e}}\right)_{2}+\left(Z_{\mathrm{f}}\right)_{2}=0.13341-0.02512 i$. The second graph shows the total impedance is minimized at $m=2$. The general trend for $h_{m}^{\prime}(k a)$ is that it rises rapidly as $m$ increases beyond $k a$. However, the third graph indicates that this increase is not drastic for $m=2$, for which $h_{2}^{\prime}(1.188)=0.12829+4.8275 i$. The logarithmic scale in the fourth graph obscures relative magnitudes, but it is evident that $\left(C_{\mathrm{e}}\right)_{2}$ is the largest coefficient.


Figure 2

As $k a$ is increased, the plots of $\left(Z_{\mathrm{e}}\right)_{m}+\left(Z_{\mathrm{f}}\right)_{m}$ and $h_{m}^{\prime}(k a)$ as functions of $m$ maintain their general shape, except that $\operatorname{Im}\left[\left(Z_{\mathrm{e}}\right)_{m}+\left(Z_{\mathrm{f}}\right)_{m}\right]$ crosses the zero axis at an increasing $m$. Figure 3 describes the behavior at $k a=4$, which does not correspond to a peak in either scattering cross section. The plot of $\operatorname{Im}\left[\left(Z_{\mathrm{e}}\right)_{m}+\right.$ $\left.\left(Z_{\mathrm{f}}\right)_{m}\right]$ crosses the zero axis between $m=12$ and $m=13$, and the minimum value of $\left|\left(Z_{\mathrm{e}}\right)_{m}+\left(Z_{\mathrm{f}}\right)_{m}\right|$ occurs at $m=11$, where it is 0.03865 . However, $h_{12}^{\prime}(4)=$ $4.4811\left(10^{-6}\right)+2.0608\left(10^{4}\right) i$, which is quite large, and this quantity is squared in the denominator of $\left(C_{\mathrm{e}}\right)_{m}$. The consequence is that the largest value of $\left(C_{\mathrm{e}}\right)_{m}$ occurs at $m=3$, but it is not especially large relative to the values for adjacent $m$. Hence, this frequency does not lead to enhanced scattering.


Figure 3

These observations suggest that the increase of $\left|h_{m}^{\prime}(k a)\right|$ for $m>k a$ is responsible for the suppression of high-frequency resonances. An example of this behavior is $k a=8.710$ at $\tilde{m}=19$, for which $\left(Z_{\mathrm{e}}\right)_{\tilde{m}}+\left(Z_{\mathrm{f}}\right)_{\tilde{m}}=8.8890\left(10^{-10}\right)-6.7896$ $\left(10^{-5}\right) i$. For comparison, at $m=18$ the value is $\left(Z_{\mathrm{e}}\right)_{18}+\left(Z_{\mathrm{f}}\right)_{18}=1.7893\left(10^{-8}\right)+$ $0.85717 i$. However, at $m=19,\left|h_{19}^{\prime}(8.71)\right|=3.2851\left(10^{4}\right)$, so the relative smallness of the impedance term is overwhelmed by the largeness of $h_{19}^{\prime}(k a)$. The largest coefficient value at $k a=8.710$ is $\left(C_{\mathrm{e}}\right)_{m}=1.4761$ at $m=7$, whereas $\left(C_{e}\right)_{19}=$ $6.5216\left(10^{-5}\right)$. Thus, the absence of high-frequency resonances in the scattering cross sections may be attributed to the fact that the spherical harmonic excitation becomes much weaker with increasing $m$, as well as the low radiation efficiency of the higher spherical harmonics of the surface displacement. (A spherical harmonic whose number is large corresponds to a short wavelength/subsonic surface motion.)

Although resonances are not present in scattering cross sections at high frequencies, the shell's flexibility does influence the scattering properties at all frequencies. As shown in Fig. 1, the value of the total scattering cross section is considerably increased by the flexibility of the shell, and the backscatter cross section shows considerable variability.

There is no need to display the cross sections for the case where the surrounding fluid is air, because they are identical to those in Figs. 12.7 and 12.8 for a rigid, stationary sphere. The flexibility of the shell has little effect on the scattering because there is an enormous difference in the properties of the shell relative to air and water. The parameter that governs the overall magnitude of $P_{\mathrm{e}}$ relative to $P_{\mathrm{b}}$ is $\rho_{0} c /\left(\rho_{\mathrm{e}} \mathrm{c}_{\mathrm{e}}\right)$, which is $3\left(10^{-5}\right)$ for air and aluminum, and 0.11 for water and aluminum. Thus, $P_{\mathrm{e}}$ is very a small contributor to the scattered field. Even resonances are insignificant relative to the blocked pressure.

### 12.7 Homework Exercises

Exercise 12.1 A solid plastic sphere immersed in an ideal liquid has density and sound speed that are slightly larger than the liquid. A plane wave is incident on this object. Specialize the Born approximation of $(r / a)\left|P_{s}\left(\bar{x}_{0}\right) / \hat{P}_{\mathrm{I}}\right|$ to fit this spherical configuration. Doing so will yield an integral representation of the scattered pressure at a field point in the farfield as a function of $k a, \rho / \rho_{0}$, and $c / c_{0}$, and the direction angles to the field point. Evaluate this integral for the case of backscatter corresponding to $\rho^{\prime} / \rho_{0}=1.02$ and $c^{\prime} / c_{0}=1.05$. Plot the backscatter amplitude in the interval $k a<2$, and compare it to the Rayleigh limit in Eq. (12.3.1).

Exercise 12.2 A submerged object is surrounded by eight equally spaced hydrophones, as shown in the sketch. Projector $A$ emits an omnidirectional signal at 20 Hz . This signal is measured to be $210 \mathrm{~dB} / / 1 \mu \mathrm{~Pa}$ at one meter from the source. The table provides the amplitude of the scattered signal measured at each hydrophone. The ambient properties for the water are $\rho_{0}=1004 \mathrm{~kg} / \mathrm{m}^{3}$ and $c=1470 \mathrm{~m} / \mathrm{s}$. A visual inspection of the object indicates that its volume is $0.12 \mathrm{~m}^{3}$. (a) Assuming that the object's physical properties are comparable to water, use the tabulated data to deduce the density and sound speed within the object. Use the Rayleigh limit of the Born approximation to analyze the data. (b) To explore the effect of measurement error, contaminate the tabulated data by adding to each value a random error that ranges over $\pm 0.4 \mathrm{~Pa}$ with a uniform probability distribution. Repeat the calculation in Part (a) using the error-contaminated data. What does the result say about the significance of measurement error?

$$
\begin{array}{l||cccccccc}
\text { Hydrophone \# } & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
\hline\left|P_{s}\right|(\mathrm{mPa}) & 1.5296 & 1.3526 & 1.1022 & 0.92514 & 0.92514 & 1.1022 & 1.3526 & 1.5296
\end{array}
$$



Exercise 12.2

Exercise 12.3 A transducer capable of transmitting and receiving is situated at point $A$, and a receiving hydrophone is situated 400 m away at point $B$. It is desired to use this arrangement to measure the physical properties of a bottle-nosed dolphin, based on the notion that it does not differ much from water. The scheme uses transducer $A$ to generate a tone burst consisting of three cycles of a 50 Hz harmonic. The scattered
signal is measured by both transducers. The received signals are plotted below. (The pulse received at position $B$ directly from the source at point $A$ has been removed.) The timescale is based on $t=0$ being the instant when the tone burst is initiated at source $A$. The measured signals allow for determination of the geometrical properties required to evaluate the Rayleigh limit of the Born approximation. It is known that the transducer $A$ generates a radially symmetric signal that is 49 kPa at 1 m from the center, and a measurement of the displacement of the dolphin indicates that its volume is $0.987 \mathrm{~m}^{3}$. The density and sound speed of water are the nominal values, $c=1480 \mathrm{~m} / \mathrm{s}$ and $\rho_{0}=1000 \mathrm{~kg} / \mathrm{m}^{3}$. Based on the assumption that the dolphin is approximately homogeneous, determine its average density and sound speed.


Exercise 12.3
Exercise 12.4 Consider the cylinder in Example 12.1 when the incident wave arrives broadside, $\psi_{\mathrm{I}}=90^{\circ}$. Determine the directivity of the scattered pressure in the plane formed by the incident wave's propagation direction and the axis of the cylinder. Frequencies of interest are $k a=0.1,1$, and 10 .

Exercise 12.5 A plane wave is incident at direction $\bar{e}_{\mathrm{I}}$ on the stationary rigid rectangular box in the sketch. The box, whose walls have negligible stiffness and mass, is filled with a liquid for which $\rho^{\prime}=0.92 \rho_{0}$ and $c^{\prime}=0.94$. The box sides are $a, b$, and $d$. (a) Use the Born approximation to derive an expression for the low-frequency scattered pressure in the form $(r / a)\left|P_{s}\left(\bar{x}_{0}\right) / \hat{P}_{\mathrm{I}}\right|$ for arbitrary $\bar{e}_{\mathrm{I}}$ and scattering direction $\bar{e}_{r}$. For the development, describe $\bar{e}_{I}$ and $\bar{e}_{r}$ in terms of spherical angles described relative to the $z$-axis. (b) Specialize the result in Part (a) to the backscatter case, $\bar{e}_{\mathrm{r}}=-\bar{e}_{\mathrm{I}}$ when a plane wave is incident in the $x z$-plane. (c) Let $\ell=\mathcal{V}^{1 / 3}$, where $\mathcal{V}$ is the volume, be a representative length scale. Compare the backscatter directivity for two shapes having the same volume: (1) $a=b=d=\ell$, (2) $a=b=0.5 \ell, d=4 \ell$.


Exercise 12.5
Exercise 12.6 The cylinder in the sketch is rigid and stationary. It is insonified by a plane wave at frequency $\omega$ whose propagation direction is parallel to the centerline of the cylinder. The radius $a$ is extremely small compared to the length $L$, and $k a \ll 1$. Derive an expression for the scattered pressure at an arbitrary field point $\bar{x}$ in the farfield.


Exercise 12.6
Exercise 12.7 The sketch shows a mass-spring oscillator consisting of a sphere whose mass is $M$ that is supported by a spring whose stiffness is $K$. A harmonic plane wave, which is traveling parallel to the direction in which the spring is oriented, causes the sphere to oscillate. The sphere is rigid, so the only motion that it executes is a rigid body translation $u=\operatorname{Re}(U \exp (i \omega t))$. It may be assumed that $k a \ll 1$. (a) Derive an expression for $U$ as a function of $k a$. (b) Derive an expression for the scattered pressure $P_{s}$ at an arbitrary field point in the farfield. (c) Determine the value of $k a$ that will maximize $\left|P_{s}\right|$.


Exercise 12.7.
Exercise 12.8 A 20 mm diameter ball bearing is insonified by a 1 kHz plane wave. Evaluate its scattering cross section in situations where it is restricted from moving,
and where it is suspended in a manner that permits it to translate. Compare the results for a ball compose of lead ( $\rho=11340 \mathrm{~kg} / \mathrm{m}^{3}, c=1322 \mathrm{~m} / \mathrm{s}$ ), to those for an aluminum ball ( $\rho=2100 \mathrm{~kg} / \mathrm{m}^{3}, c=6410 \mathrm{~m} / \mathrm{s}$ ).

Exercise 12.9 A plane wave at 60 Hz in air is obliquely incident on a stationary rigid disk whose radius is 200 mm . Its direction of propagation is $20^{\circ}$ off the normal to the disk. Determine the scattering cross section, the backscatter cross section, and the target strength corresponding to this angle of incidence.

Exercise 12.10 A land mine is buried in sand, $\rho_{0}=1140 \mathrm{~kg} / \mathrm{m}^{3}, c=105 \mathrm{~m} / \mathrm{s}$, which may be regarded as a liquid. It is insonified by a 15 Hz plane wave whose amplitude is 500 Pa . The mine is buried sufficiently deeply that reflections of the scattered wave from the surface may be ignored. It is a $200-\mathrm{mm}$-long cylinder whose radius is 150 mm , and both ends are closed with hemispherical caps. The mass of the cylinder is 120 kg and the virtual mass coefficients are $\Lambda_{1,1}=\Lambda_{2,2}=10 \mathcal{V}$, $\Lambda_{3,3}=2.5 \mathcal{V}$. Determine its scattering cross section and target strength in the situation where $\bar{e}_{\mathrm{I}}=-\bar{e}_{x}$. Do these values depend strongly on whether the mine is stationary or movable?

Exercise 12.11 A plane harmonic wave in the atmosphere is incident on an object that is symmetric about the $z$-axis, with $\bar{e}_{\mathrm{I}}=-\bar{e}_{z}$. The frequency is 500 Hz . The scattered pressure at a point in the farfield is described in terms of spherical harmonics as

$$
P_{\mathrm{s}}=P_{\mathrm{I}} \frac{e^{-i k r_{0}}}{k r_{0}} \sum_{m=0}^{\infty} B_{m} P_{m}\left(\cos \psi_{0}\right) e^{i(m+1) \pi / 2}
$$

The coefficient values are $B_{0}=0.5, B_{1}=0.2-0.4 i, B_{2}=1.5+0.3 i, B_{3}=-$ $0.2+0.2 i \mathrm{~Pa}$, while all higher values are zero. The forward scattered pressure is 2 Pa at $r=10 \mathrm{~m}$. (a) What is the backscattered pressure at $r=10 \mathrm{~m}$. (b) What are the differential scattering cross section values at $\psi=0,90^{\circ}$, and $180^{\circ}$ ? (c) What is the scattering cross section of this object? (d) What is the target strength?

Exercise 12.12 The axisymmetric object in the sketch is a body of revolution. It is generated by rotating about the centerline a circular arc of radius $4 a$ whose center is situated at $5 a$ from the centerline. A plane harmonic wave traveling in the direction $-\bar{e}_{x}$ is incident on the object. According to geometrical acoustics, what is the backscatter cross section of this object? Would this value change if the incident wave's direction deviated slightly from broadside incidence?


Exercise 12.12
Exercise 12.13 A crude model for scattering from a rough surface considers normal incidence of a plane wave on a surface that is nominally flat, but actually has a sinusoidal variation in its elevation. Such a situation is depicted in the sketch, where the elevation is $h=h_{0} \cos (2 \pi x / L)$. If the acoustic wavelength $2 \pi / k$ is much smaller than the spatial period $L$ of the surface, it is permissible to invoke geometrical acoustics. Construct a sketch of the manner in which the incident rays scatter. What general conclusions can be drawn from this construction?


Exercise 12.13
Exercise 12.14 A plane harmonic wave propagating horizontally in water is incident on a solid sphere. The sphere is suspended by a cable that permits it to move in the horizontal plane. Derive expressions for the backscattered pressure and the velocity of the center of the sphere as a function of $k a<10$ for the case where the sphere's mass is $5 \pi a^{3} \rho_{0}$.

Exercise 12.15 Exercise 12.7 entailed an analysis of scattering from a springsupported rigid sphere in the low-frequency regime, $k a \ll 1$. (a) Perform the investigation requested there, but do not assume that $k a$ is small. (b) Evaluate $\rho_{0} c\left|V_{C}\right| /\left|P_{\mathrm{I}}\right|$ as a function of $k a<5$ for four parametric combinations: $M=\sigma \rho_{0} \pi a^{3}$ and $K=$ $\chi M(c / a)^{2}$ with $\sigma=2$ and $8, \chi=1$ and 2 . Identify the frequency at which $\left|V_{C}\right|$ is maximized.

Exercise 12.16 A 3-m diameter steel shell is neutrally buoyant in a very large vat of ethyl alcohol ( $\rho_{0}=785 \mathrm{~kg} / \mathrm{m}^{3}, c=1144 \mathrm{~m} / \mathrm{s}$ ). It is insonified by a plane harmonic wave. Determine the scattering directivity, that is, $(r / a)\left|P_{s} / P_{\mathrm{I}}\right|$ as a function of $\psi$.

Identify the portion of this function that is attributable to the elasticity of the shell. Carry out this evaluation for cases where the frequency of this wave is $99 \%$ of the upper natural frequency of the $m=0,1$, and 2 in-vacuo modes.

Exercise 12.17 The sketch shows a plane wave that is incident on a rigid hemisphere mounted on an infinite rigid baffle. The sum of the incident and scattered fields must have zero particle velocity normal to the baffle, as well as the normal to the hemisphere's surface, that is, $\left(\bar{v}_{\mathrm{I}}+\bar{v}_{s}\right) \cdot \bar{e}_{z}=0$ for $r>a, \psi=\pi / 2$ and $\left(\bar{v}_{\mathrm{I}}+\bar{v}_{s}\right) \cdot$ $\bar{e}_{r}=0$ for $r=a, 0 \leq \psi<\pi / 2$. A suitable representation draws on the development in Sect. 7.2.3, which showed that the field radiated by a vibrating hemisphere on a rigid baffle can be represented as a series of the even-numbered spherical harmonics. At the same time, if the hemisphere was not present, the incident wave would be reflected with the reflection coefficient being one. Thus, an ansatz for the field scattered by the hemisphere and baffle is

$$
\begin{gathered}
P=P_{\mathrm{I}}+P_{R}+P_{h} \\
P_{\mathrm{I}}=\hat{P}_{\mathrm{I}} e^{+i k z}, \quad P_{R}=\hat{P}_{\mathrm{I}} e^{-i k z}, \quad P_{h}=\sum_{m=0}^{\infty} B_{m} h_{2 m}(k r) P_{2 m}(\cos \psi)
\end{gathered}
$$

Derive an expression that describes the $B_{m}$ coefficients.


Exercise 12.17
Exercise 12.18 Consider the hemisphere-baffle configuration in Exercise 12.17. The scattered pressure is $P_{s}=P_{R}+P_{h}$. The frequency of interest is $k a=10$. (a) Evaluate the scattered pressure as a function of distance along the $z$-axis. Compare the result to the geometrical acoustics approximation for the backscatter. (b) Evaluate the dependence of the scattered pressure along a line parallel to the baffle at $z=4 a$. (c) Evaluate the dependence of the scattered pressure on the polar angle $\psi$ at $r=4 a$.

## Chapter 13 Nonlinear Acoustic Waves

When told that a system's response is quite large, a knowledgeable individual will anticipate that nonlinear effects are important. This certainly is the case in acoustics, as exemplified by an explosion, wherein peak pressures might be well above atmospheric. However, even if the pressure is not large, there is another parameter whose largeness might require consideration of nonlinear effects. That parameter is the propagation distance. The signal radiated by a transducer might not seem to be especially large, but nonlinear effects grow with increasing distance. We might observe harmonic and intermodulation distortion of the waveform, decreased signal strength, or the formation of shocks (discontinuous changes of the waveform). Although these effects would seem to be deleterious, there are some applications that rely on them.

Analysis of nonlinear effects encounters two fundamental difficulties relative to linear theory. The most obvious is the inapplicability of the principle of superposition. But there is another. In our studies thus far, conversion to the frequency domain made it possible to analyze many systems. The concept of using a complex exponential to remove the explicit dependence on time ceases to be valid in the presence of nonlinearity. The reason becomes evident when we consider a nonlinear term such as $p^{2}$. If $p$ oscillates at frequency $\omega$, then $p^{2}$ contains a term whose frequency is $2 \omega$. The occurrence of such a term in the field equation for $p$ means that $p$ cannot be a pure harmonic. The consequence is that we will need to work in the time domain, so far fewer systems are amenable to analysis.

In many treatises, the term finite amplitude wave is used to refer to a signal whose amplitude is sufficiently large that nonlinear effects cannot be ignored. The notion behind this terminology is that ignoring terms like $p^{2}$ is equivalent to considering $p$ to be infinitesimal. Whether it is allowable to do so depends on several factors, most notable of which are the amplitude of the signal and the propagation distance. If either is large, one might need to consider nonlinear effects.

We will not pursue a historical survey of the numerous approaches that have been used. Rather, the methods we shall explore have been selected either for their
generality or for their directness. We will follow Riemann's seminal analysis, ${ }^{1}$ which is based on the fundamental properties of differential equations. One of its beneficial attributes is that it does not require specification of the excitation. Another approach we will develop applies Fourier series to analyze systems in which the excitation is temporally periodic. The idea is not to convert the problem to the frequency domain. Rather, the formulation yields differential equations for the Fourier coefficients that are easier to solve. A third approach for our consideration is perturbation analysis, which decomposes the nonlinear problem to a sequence of linearized problems. This method cannot proceed if the analogous linear problem cannot be solved.

### 13.1 Riemann's Solution for Plane Waves

It is remarkable that Riemann's solution was published in 1859, and that it was contemporaneous with Earnshaw's analysis ${ }^{2}$ of the same problem. The two papers followed very different paths. Riemann used the continuity and momentum principles to obtain equations governing the fundamental state variables: pressure, particle velocity, and density. Earnshaw formulated the analysis in terms of displacement of a specific particle in what is called a Lagrangian kinematical description. Most importantly for us, Riemann's analysis proceeds in a logical manner, whereas Earnshaw's requires a degree of intuition. The results of the analyses are equivalent. Both may be found in treatises by Rayleigh ${ }^{3}$ and Lamb. ${ }^{4}$ Poisson ${ }^{5}$ actually was the first to solve this problem, so some refer to the final result as the Poisson solution. However, that treatment assumed isothermal conditions. We will follow Riemann's analysis because of its directness and generality.

### 13.1.1 Analysis

The fluctuating acoustic part of the pressure is $p$, so the absolute pressure in the current state is $p_{0}+p$. The corresponding density is $\rho$. The direction of propagation is designated as $x$. The particle velocity, positive in the sense of increasing $x$, is $v$. The basic equations of continuity and momentum for a one-dimensional wave are

[^66]Eqs. (2.1.6) and (2.1.10). They are

$$
\begin{gather*}
\frac{\partial \rho}{\partial t}+\frac{\partial}{\partial x}(\rho v)=0 \\
\rho\left(\frac{\partial v}{\partial t}+v \frac{\partial v}{\partial x}\right)=-\frac{\partial p}{\partial x} \tag{13.1.1}
\end{gather*}
$$

In addition, we retain the assumption of adiabatic compression and expansion, based on the expectation that the pressure varies at a rate that is too high to permit significant heat transfer.

The notion motivating the Riemann solution is that there is a family of curves in space-time along which $p$ and $v$ are constant. If this is true, then as long as we stay on one of these curves, we can consider $p$ to be a function of $t$ and $x$, and $v$ to be a function of $p$. A curve $t$ as a function of $x$ (or $x$ as a function of $t$ ) having this property is a characteristic. Thus, along a characteristic, the functional dependence is $v=v(p)$. As a consequence of the adiabatic assumption, the equation of state may be taken to be $\rho=\rho(p)$.

The task is to solve the first-order conservation equations for the relation between $x$ and $t$ consistent with the assumed functional dependence on $p$. The derivatives of $\rho$ and $v$ appearing in the basic equations are converted by the chain rule to

$$
\begin{gather*}
\frac{\partial \rho}{\partial t}=\frac{1}{d p / d \rho} \frac{\partial p}{\partial t}, \quad \frac{\partial}{\partial x}(\rho v)=\frac{d}{d p}(\rho v) \frac{\partial p}{\partial x}  \tag{13.1.2}\\
\frac{\partial v}{\partial t}=\frac{d v}{d p} \frac{\partial p}{\partial t}, \quad \frac{\partial v}{\partial x}=\frac{d v}{d p} \frac{\partial p}{\partial x}
\end{gather*}
$$

In linear theory, the square of the sound speed is defined to be $d p / d \rho$ in the ambient state. In the nonlinear formulation, the relevant value is this derivative at the current state. Many references define $c^{2}$ to be this value and denote $c_{0}$ as the linear speed of sound. It is less confusing if $c$ is retained as the linear speed of sound, so $\tilde{c}^{2}$ is $d p / d \rho$ at the current state,

$$
\begin{equation*}
\tilde{c}=\left(\frac{d p}{d \rho}\right)^{1 / 2} \equiv\left(\frac{1}{d \rho / d p}\right)^{1 / 2} \tag{13.1.3}
\end{equation*}
$$

When Eq.(13.1.2) is substituted into the basic continuity and momentum equations, the result is

$$
\begin{align*}
& \frac{1}{\tilde{c}^{2}} \frac{\partial p}{\partial t}+\left(\frac{1}{\tilde{c}^{2}} v+\rho \frac{d v}{d p}\right) \frac{\partial p}{\partial x}=0 \\
& \rho \frac{d v}{d p} \frac{\partial p}{\partial t}+\left(\rho v \frac{d v}{d p}+1\right) \frac{\partial p}{\partial x}=0 \tag{13.1.4}
\end{align*}
$$

Because of the assumed interdependence of variables, the only quantities that depend explicitly on $x$ and $t$ are the derivatives of $p$. Thus, the preceding may be considered to be a pair of first-order equations for $\partial p / \partial t$ and $\partial p / \partial x$ in which the coefficients are functions of $p$. The matrix form of the equations is

$$
\left[\begin{array}{c}
\frac{1}{\tilde{c}^{2}}\left(\frac{1}{\tilde{c}^{2}} v+\rho \frac{d v}{d p}\right)  \tag{13.1.5}\\
\rho \frac{d v}{d p}\left(\rho v \frac{d v}{d p}+1\right)
\end{array}\right]\left\{\begin{array}{l}
\partial p / \partial t \\
\partial p / \partial x
\end{array}\right\}=\left\{\begin{array}{l}
0 \\
0
\end{array}\right\}
$$

These equations are homogeneous. If the relation between $p$ and $v$ were arbitrary, their solution would be that both $\partial p / \partial t$ and $\partial p / \partial x$ are zero. These conditions lead to the trivial solution that $p$ is constant. A nontrivial solution is obtained only if the pair of equations cannot be solved. In other words, a nontrivial solution along a characteristic curve exists only if the determinant of the coefficient matrix vanishes. This condition reduces to

$$
\begin{equation*}
\frac{1}{\tilde{c}^{2}}-\left(\rho \frac{d v}{d p}\right)^{2}=0 \tag{13.1.6}
\end{equation*}
$$

Thus, we find that

$$
\begin{equation*}
\frac{d v}{d p}= \pm \frac{1}{\rho \tilde{c}} \tag{13.1.7}
\end{equation*}
$$

The equation of state relates $\rho$ and $p$, and $\tilde{c}$ is obtained by differentiating that relation. Hence, the preceding constitutes a first-order ordinary differential equation that relates $p$ and $v$ on a characteristic. It may be integrated, subject to the condition that the fluid is quiescent in the ambient state, $v=0$ if $p=0$. Thus, the integrated relation is

$$
\begin{equation*}
v=\int_{0}^{p} \frac{d p^{\prime}}{\rho\left(p^{\prime}\right) \tilde{c}\left(p^{\prime}\right)} \tag{13.1.8}
\end{equation*}
$$

We shall proceed on the assumption that we have evaluated this integral.
It is useful at this juncture to recall the relation between the particle velocity and the pressure fluctuation in a linear plane wave, $\bar{v}=\bar{e} p /\left(\rho_{0} c\right)$. The propagation here may be in either the positive or negative $x$ direction, so $\bar{e}= \pm \bar{e}_{x}$, and $\bar{v}=v \bar{e}_{x}$. Thus, if we were to differentiate the relation for a linear wave, we would obtain $d v / d p=$ $\pm 1 /\left(\rho_{0} c\right)$. The present nonlinear analysis tells us that the effective characteristic impedance is the value at the current state. An analogous relationship occurs when the force in a spring depends nonlinearly on the strain, which is equivalent to the strain being a function of the force. Then, the incremental force $d F$ required to increase the strain by $d \varepsilon$ depends on the value of the current force. In other words, $d F=K(F) d \varepsilon$, where $K=1(d \varepsilon / d F)$. For a linear spring, $K$ is the (constant) stiffness, so $K(F)$ is
said to be the tangential stiffness. If we were to use similar terminology for a nonlinear plane wave, we would call $\rho \tilde{c}$ the tangential characteristic impedance of the fluid.

A corollary of the coefficient matrix in Eq. (13.1.5) having a zero determinant is that the two differential equations are equivalent. When Eq. (13.1.7) applies, both require that

$$
\begin{equation*}
\frac{\partial p}{\partial t}+(v \pm \tilde{c}) \frac{\partial p}{\partial x}=0 \tag{13.1.9}
\end{equation*}
$$

Note that the correspondence of alternative signs must be maintained, so a negative sign corresponds to a wave that propagates in the direction of decreasing $x$.

The preceding tells us that anywhere along a characteristic, $\partial p / \partial t$ is proportional to $\partial p / \partial x$, where the proportionality factor is $-(v \pm \tilde{c})$, whose constancy follows from the fact that both $v$ and $\tilde{c}$ depend on the $p$ value for that characteristic. Constancy of $p$ along a characteristic may be expressed alternatively by equating $p$ at any two ( $x, t$ ) pairs on the same characteristic. The points we select are differentially apart, so we require that $p(x+d x, t+d t)=p(x, t)$. Application of a Taylor series leads to

$$
\begin{equation*}
\left.\left(\frac{\partial p}{\partial x} d x+\frac{\partial p}{\partial t} d t\right)\right|_{\operatorname{constant} p(x, t)}=0 \tag{13.1.10}
\end{equation*}
$$

When we substitute Eq. (13.1.7), the result is

$$
\begin{equation*}
\left.\left[\frac{d x}{d t}-(v \pm \tilde{c})\right]\right|_{\text {constant } p(x, t)}=0 \tag{13.1.11}
\end{equation*}
$$

Because $v$ and $\tilde{c}$ are constant in this condition, it follows that $d x / d t$ is constant. In other words, the slope $d x / d t$ defining where a specific $p$ occurs in space-time is constant at $\pm \tilde{c}+v$. Stated differently, the characteristics are straight lines! If $\tilde{p}$ is known at position $x_{0}$ at time $t_{0}$, then the same pressure will occur at position $x$ at time $t$, with

$$
\begin{equation*}
x-x_{0}=(v \pm \tilde{c})\left(t-t_{0}\right) \tag{13.1.12}
\end{equation*}
$$

This simple property has some of profound implications, which we shall now examine.

### 13.1.2 Interpretation

From a viewpoint of phenomenology, the most important aspect is that the phase velocity of a plane wave that propagates in the direction of increasing $x$ is $(\tilde{c}+v) \bar{e}_{x}$. The difference between $\tilde{c}$ and the linear speed of sound, $c$, is a consequence of the variability of the equation of state. Consider an ambient state in which the fluid is at rest and the pressure is the sum of the actual ambient pressure $p_{0}$ plus the acoustic
perturbation $p$. Then, $\tilde{c}$ would be the linear sound speed $c$. The phase speed also differs from $c$ because the particle velocity is nonzero. Suppose that in this ambient state, the fluid flows uniformly at $v \bar{e}_{x}$. In this condition, the fluid is at rest relative to a reference frame that translates in unison with the ambient flow. That reference frame is inertial, so sound propagates relative to the moving reference frame at speed $\tilde{c} \bar{e}_{x}$. The propagation velocity as seen by a fixed observer is the velocity relative to the moving reference frame plus the velocity of the reference frame, that is, $\tilde{c} \bar{e}_{x}+v \bar{e}_{x}$. In the terminology of nonlinear mechanics, the difference between $\tilde{c}$ and $c$ is said to be a consequence of material nonlinearity. The fact that $v$ adds to the speed of sound is said to be a convective nonlinearity. Although the two sources of nonlinearity are unrelated in regard to their cause, they affect the propagation speed similarly.

The (linear) d'Alembert solution for plane waves actually is quite similar to the present result. It states that constant values of $p$ propagate along lines for which either $t-x / c$ or $t+x / c$ is constant. These lines in space-time are the characteristics of the linear wave equation. Figure 2.3 describes the characteristics as graphs of $t$ as a function of $x$. Similar graphs are extremely useful for nonlinear waves. Let us begin with a boundary value problem, in which $v$ at $x=0$ is a specified function of $t$.

The propagation is taken to be in the sense of increasing $x$, so the positive sign preceding $\tilde{c}$ applies in Eq. (13.1.11) and everywhere else. The abscissa is $x$, so the slope of a characteristic line is $d t / d x=1 /(\tilde{c}+v)$. This slope is less than $1 / c$ if $p>0$. To see why, note that Eq. (13.1.7) in this case gives $d v / d p>0$, which means that $v>0$ if $p>0$. Furthermore, the equation of state for ideal gases and liquids is such that $\tilde{c}>c$ if $p>0$. Because the slope is less than $1 / c$, a characteristic along which $v>0$ describes propagation that is faster than the linear speed of sound. The converse is true for characteristics associated with $v<0$. They describe propagation at a speed that is less than $c$. (It is best to avoid using terms like subsonic or supersonic here.)

To visualize these properties, let us suppose we know the particle velocity at $x=0$ as a function of time. (This is a simplification of the actual boundary condition for a source, such as a vibrating wall. The proper description of this condition may be found in the next section.) The case of a single period of a sine pulse starting at $t=0$ is plotted along the $t$-axis in Fig.13.1.

Fig. 13.1 Characteristics lines for propagation of the nonlinear wave generated by a sinusoid velocity pulse at $x=0$


The characteristics through the instants when $v=0$ at $x=0$ represent propagation at the linear speed of sound because $v$ and $p$ are zero along them. Because $v$ at the boundary is zero for $t<0$ and $t>2 \pi / \omega$, the pressure is zero in the region below the lowest characteristic and above the uppermost characteristic. Signals for which $v>0$ cover the same distance in a shorter elapsed time, whereas signals for which $v<0$ take longer to travel a given distance. It follows that the shape of a waveform will change as the wave propagates. A linear wave, which is described by the d'Alembert solution, is said to be dispersionless, because all waveforms maintain their shape. In a waveguide, nonplanar modes have a phase speed that depends on the frequency, which means that the waveforms undergo frequency dispersion. The changing shape of the waveform for a nonlinear plane wave is said to be amplitude dispersion.

Because the characteristics for different $v$ values are not parallel, it is evident that these lines will eventually intersect. At such intersections, the analysis indicates that two signals having different values of $v$ will exist at the same $x$ and $t$. This is an impossible condition. What actually happens is that discontinuities, which are called shocks, form. We will explore this phenomenon later. For now, we will consider the intersection of characteristics as a limit beyond which the Riemann solution is not valid.

A consequence of $p$ being constant on a characteristic is that lines that originate at instants when $p$ is maximum will retain this property. Similarly, lines that originate from minimum or zero $p$ remain so. We could use this attribute to construct a sketch of a waveform at fixed $x$ or a profile at fixed $t$. This is one approach described in the next section. However, before we address the quantitative evaluation of a signal, let us consolidate what we have learned.

The pressure and particle velocity are constant along a characteristic, and a characteristic is the locus of points in space-time at which $t-x /(\tilde{c}+v)$ is a specified constant. This constant value is a phase variable for the wave. We shall use $\tau$ to denote it, so that

$$
\begin{equation*}
p=F(\tau), \tau=t-\frac{x}{(\tilde{c}+v)} \tag{13.1.13}
\end{equation*}
$$

The logical question at this juncture is what is $F(\tau)$ ? The answer lies in matching the preceding general form to the manner in which the signal is generated.

### 13.1.3 Boundary and Initial Conditions

Consider the situation where the plane wave is generated by movement of a large piston or a wall that translates as a rigid body. Let $x=X(t)$ denote the (known) position of the wall at any instant. Then, continuity of particle velocity requires that the velocity of the fluid at the current location of the piston face equals the piston's velocity, that is,

$$
\begin{equation*}
v=\dot{X}(t) \text { at } x=X(t) \tag{13.1.14}
\end{equation*}
$$

This is referred to as a moving boundary condition. Because of the intricate manner in which the state variables are related, we shall take a somewhat reverse approach, in which relations that we seek are taken to be known.

Because the boundary condition imposes a particle velocity, the first step is to remove pressure from the functional dependencies. Toward that end, we invert the expression of $v$ in terms of $p$ obtained from Eq.(13.1.8). This step allows us to consider the relationship to be $p=p(v)$. The implication of this switch of dependencies is that $\tilde{c}$ is known if we know $v$. This feature is made explicit by letting $\tilde{C}$ be the value of $\tilde{c}$ at a specified $v$, so that

$$
\begin{equation*}
\tilde{C}(v)=\tilde{c}(p(v)) \tag{13.1.15}
\end{equation*}
$$

It follows that the slope of a characteristic is $d x / d t=\tilde{C}(v)+v$.
Given this, let us track a signal that departs from the boundary at time $\tau$. The particle velocity everywhere along the characteristic for this signal is $\dot{X}(\tau)$, so its slope is $\tilde{C}(\dot{X}(\tau))+\dot{X}(\tau)$. Also, the starting point of the characteristic is $(X(\tau), \tau)$. The time required for this signal to arrive at an arbitrary point $(x, t)$ is $t-\tau$, and the propagation distance to this point is $x-X(\tau)$. Equating $[x-X(\tau)] /(t-\tau)$ to $d x / d t$ yields

$$
\begin{equation*}
\frac{x-X(\tau)}{t-\tau}=\tilde{C}(\dot{X}(\tau))+\dot{X}(\tau) \tag{13.1.16}
\end{equation*}
$$

A minor rearrangement of terms yields a solution whose form is like the one derived by Earnshaw, specifically

$$
\begin{equation*}
\tau=t-\frac{x-X(\tau)}{\tilde{C}(\dot{X}(\tau))+\dot{X}(\tau)}, \quad v=\dot{X}(\tau) \tag{13.1.17}
\end{equation*}
$$

These relations fully describe the signal. To realize this, suppose we select a value of $\tau$. Given a function $X(t)$, we may evaluate $X$ and $\dot{X}$ at $t=\tau$, which sets the value of $\tilde{C}(\dot{X}(\tau))$. Then, the first of the above relations gives the locus of $x, t$ values for passage of this signal, that is, it defines the characteristic along which the particle velocity is $\dot{X}(\tau)$. The value of $p$ along this characteristic may be found from the relation $p=p(v)$ that we would determine from Eq.(13.1.8). Note that the derived solution is consistent with Eq.(13.1.13), because it merely entails measuring the propagation distance from the location of the piston, rather than the reference location at which $x=0$.

Despite the availability of the general solution in Eq.(13.1.17), it seldom is important to satisfy the moving boundary condition. To see why this is so, consider Fig. 13.2, which depicts the position of the wall in space-time as the curve $x=X(\tau)$ that oscillates about the $t$-axis. The specific function is one cycle of $X=-(\varepsilon c / \omega) \cos \omega \tau)$, which leads to $\dot{X}=\varepsilon c \sin (\omega \tau)$. The parameter $\varepsilon$ is the acoustic Mach number. Later examination will show that $\varepsilon$ is a very small number, with values exceeding 0.001 being extraordinarily uncommon.


Fig. 13.2 Characteristics originating from the instantaneous position $X(\tau)$ of a moving boundary whose nominal position is $x=0$. The solid lines are the result of satisfying velocity continuity at the instantaneous position of the boundary, whereas the dashed lines are the characteristics if the boundary's velocity is approximated as occurring at $x=0$

Each point on the $x=X(\tau)$ curve is the starting point for a characteristic, where $\tau$ is the instant at which the associated signal was generated. The slope of any characteristic is $\Delta x / \Delta t=\tilde{C}(\dot{X}(\tau))+\dot{X}(\tau)$. The characteristics described in Fig. 13.2 are those for values of $\tau$ at which $\dot{X}$ is either zero, or a maximum, or a minimum. These characteristics are compared in the figure to the nominal characteristic that would result if the value of $\dot{X}(\tau)$ occurred at $x=0$. For the selected $X$ function, the maximum and minimum values of $\dot{X}$ occur when $X=0$. These characteristics are the same as the corresponding nominal ones. This arrangement is contrasted with those associated with instants when $X$ is a maximum or minimum. The value of $\dot{X}$ is zero at these instants, which means that the phase speed is $c$, but these characteristics are shifted by the maximum amount relative to the nominal value.

The key aspect is that satisfying the moving boundary condition shifts a characteristic parallel to the $x$-axis by $X(\tau)$. This effect is quite different from the effect of the dependence of phase speed on particle velocity, which changes the slope of the characteristics. Consequently, the associated time shift increases proportionally to $x$. In the vicinity of the boundary, these time shifts might be comparable. However, with increasing $x$, the velocity dependence of the phase speed overwhelms the moving boundary condition effect.

It follows that if a signal's amplitude is within the range of common experience, the error resulting from applying velocity continuity at the undisplaced position of a boundary is insignificant. This allows us to replace the actual boundary condition in Eq. (13.1.14) with an approximate stationary boundary condition that sets $v=\dot{X}(t)$ at $x=0$. Correspondingly, the particle velocity anywhere is described by removing $X(\tau)$ from the numerator of Eq. (13.1.17). Thus, the solution reduces to

$$
\begin{equation*}
v=\dot{X}(\tau), \quad \tau=t-\frac{x}{\tilde{C}(v)+v} \tag{13.1.18}
\end{equation*}
$$

As in the case of the moving boundary formulation, the corresponding value of $p$ may be found by inverting Eq. (13.1.18). This relation requires specification of the equation of state, which is the topic for the next section.

Figure 13.1 describes the signal that propagates in the positive $x$ direction when the excitation at $x=0$ is a known function of time. How would the analysis be altered if the domain was $x<0$ ? In that case, the negative sign in Eqs. (13.1.7) and (13.1.11) would apply. Hence, the slope of a characteristic would be $d t / d x=-1 /(\tilde{c}-v)$. As before $p>0$ corresponds to $\tilde{c}>c$. However, now $d v / d \rho<0$, so $p>0$ corresponds to $v<0$. Thus, as was true for the previous case, the phase speed of a positive pressure is $\tilde{c}+|v|>c$, and the phase speed of a negative pressure is $\tilde{c}-|v|<c$.

When we studied linear wave propagation, we satisfied arbitrary initial conditions by combining the d'Alembert solution for waves that propagate in either direction. This process cannot be implemented for nonlinear waves. A general reason is that solutions cannot be superposed. If we add solutions for forward and backward propagation, the result will generate product terms that do not fit the general scheme. A more specific reason is that each characteristic line describes a constant pressure value. If we contemplate combining waves that propagate in both directions, characteristic curves for propagation in each direction would intersect. The pressure at any ( $x, t$ ) would be the sum of the value for the forward and backward characteristics that intersect there. Thus, any attempt to combine forward and backward waves would contradict the condition that the pressure is constant along a characteristic.

The consequence is that we may only use the Riemann solution to satisfy a limited initial value problem. Suppose we know that a forward propagating wave has been established somehow. (The argument is equally applicable to a backward propagating wave.) Further suppose that at some instant, which we designate as $t=0$, we have measured the profile $v$ as a function of $x$. This situation is depicted in Fig. 13.3. Other than where they originate from, this diagram shows that there is no difference between the characteristics for this initial value problem and those in the previous figure for a boundary value problem.

Fig. 13.3 Characteristics for a nonlinear wave that propagates in the sense of increasing $x$. The properties of this wave are taken to be known at $t=0$


The task of actually satisfying this type of initial value problem is expedited by a minor modification. The phase variable for a boundary value problem, Eq. (13.1.13),
is $\tau=t-x /(\tilde{c}+v)$. This definition facilitates matching the general solution to whatever function is specified at the boundary. For the initial value problem, we wish to match the general solution to whatever $x$ function is specified at $t=0$. This process is assisted by redefining the phase variable to be $\theta=x-(\tilde{c}+v) t$, so that $\theta=x$ at $t=0$. The corresponding form of the general solution is

$$
\begin{equation*}
p=f(\theta), \quad \theta=x-(\tilde{c}+v) t \tag{13.1.19}
\end{equation*}
$$

This is the form obtained by Poisson. ${ }^{6}$

### 13.1.4 Equations of State

If we wish to construct waveforms ( $p$ as a function of $t$ at specified $x$ ) and profiles ( $p$ as a function of $x$ at specified $t$ ), it is necessary to evaluate the integral in Eq. (13.1.8), which relates the values of $p$ and $v$ along a characteristic. Doing so requires specification of the equation of state, which relates $p$ and $\rho$. We begin with the case of an ideal gas that is compressed and expanded adiabatically,

$$
\begin{equation*}
\frac{p}{p_{0}}+1=\left(\frac{\rho}{\rho_{0}}\right)^{\gamma} \Longleftrightarrow \frac{\rho}{\rho_{0}}=\left(\frac{p}{p_{0}}+1\right)^{1 / \gamma} \tag{13.1.20}
\end{equation*}
$$

The linear speed of sound in an ideal gas is $c^{2}=\gamma p_{0} / \rho_{0}$. We use this relation to eliminate the ambient pressure, which converts the preceding relations to

$$
\begin{equation*}
\frac{p}{\rho_{0} c^{2}}=\frac{1}{\gamma}\left(\frac{\rho}{\rho_{0}}\right)^{\gamma}-1 \Longleftrightarrow \frac{\rho}{\rho_{0}}=\left(\frac{\gamma p}{\rho_{0} c^{2}}+1\right)^{1 / \gamma} \tag{13.1.21}
\end{equation*}
$$

From this, we find that

$$
\begin{equation*}
\frac{\tilde{c}}{c}=\left(\frac{\gamma p}{\rho_{0} c^{2}}+1\right)^{\frac{\gamma-1}{2 \gamma}} \tag{13.1.22}
\end{equation*}
$$

To determine the relation between $p$ and $v$, we substitute $\rho$ from Eq.(13.1.21) and $\tilde{c}$ from the preceding into Eq. (13.1.8), which becomes

$$
\begin{equation*}
v=\frac{1}{\rho_{0} c} \int_{0}^{p}\left(\frac{\gamma p^{\prime}}{\rho_{0} c^{2}}+1\right)^{-\left(\frac{\gamma+1}{2 \gamma}\right)}\left(\frac{\rho_{0} c^{2}}{\gamma}\right) d\left(\frac{\gamma p^{\prime}}{\rho_{0} c^{2}}\right) \tag{13.1.23}
\end{equation*}
$$

[^67]The result is

$$
\begin{equation*}
v=c\left(\frac{2}{\gamma-1}\right)\left[\left(\frac{\gamma p}{\rho_{0} c^{2}}+1\right)^{\left(\frac{\gamma-1}{2 \gamma}\right)}-1\right] \tag{13.1.24}
\end{equation*}
$$

This equation may be solved for $p$ as a function of $v$, which is

$$
\begin{equation*}
\frac{p}{\rho_{0} c^{2}}=\frac{1}{\gamma}\left[\left(1+\frac{\gamma-1}{2} \frac{v}{c}\right)^{2 \gamma /(\gamma-1)}-1\right] \tag{13.1.25}
\end{equation*}
$$

We can gain some insight into the significance of nonlinearity by examining graphs of $\tilde{c}$ and $v$ as functions of $p$; Fig. 13.4 describes both functions nondimensionally for the case where the gas is air, $\gamma=1.4$. The dotted lines are the relations for linear acoustics, $\rho=\rho_{0}+p / c^{2}, \quad \tilde{c}=c, \quad v=p /\left(\rho_{0} c\right)$.

Fig. 13.4 Nonlinear dependence of $\rho, \tilde{c}$, and $v$ on the pressure in an ideal gas for $\gamma=1.4$. The analogous relation for a linearized equation of state is shown as a dotted line


The linear acoustic approximations are very close to the corresponding exact curve in the range $|p| / p_{0}<0.2$. This range corresponds to an acoustic amplitude that is $20 \%$ of the ambient pressure. In air, for which $p_{0}$ is 1 atm , this amplitude is approximately 20 kPa , which corresponds to a harmonic signal at $217 \mathrm{~dB} / / 20 \mu \mathrm{~Pa}$. Such a signal is beyond anything that would be encountered in a conventional acoustic application.

The fact that these relations are close to those of linear theory does not mean that nonlinearity is unimportant, because the phase variable $\tau$ is not simply $t-x / c$. One could say that the linearized analysis gives the correct solution, except that it places that solution in the wrong space-time location.

The logical question at this juncture is: What modifications are required if the fluid is a liquid? The equation of state for an adiabatic process in a liquid is given by Eq. (2.3.18),

$$
\begin{equation*}
\frac{p}{\rho_{0} c^{2}}=\left(\frac{\rho}{\rho_{0}}-1\right)+\frac{1}{2} \frac{B}{A}\left(\frac{\rho}{\rho_{0}}-1\right)^{2}+\cdots \tag{13.1.26}
\end{equation*}
$$

This series relation is derived experimentally. The value of $A$ is the bulk modulus, $A=K=\rho_{0} c^{2}$. It is convenient to consider $B / A$ to be a single coefficient, because that is the manner in which $B$ arises in nonlinear waves. The value of $B / A$ for water ranges from 4.2 for distilled water at atmospheric pressure and $0^{\circ} \mathrm{C}$ to 6.2 for distilled water between 200 and 4000 atm at $30^{\circ} \mathrm{C}$. A nominal value $B / A=5$ may be used if environmental conditions are unspecified. It also is the value for seawater at atmospheric pressure, $20^{\circ} \mathrm{C}$, and 35 parts per thousand salinity. ${ }^{7}$

The equation of state for a liquid contains only two terms because the bulk modulus is very large. For example, $\rho_{0} c^{2}=2.2 \mathrm{GPa}$ for water. Hence, large values of $\rho / \rho_{0}-1$ lead to very large values of $p$ that are outside the usual realm of acoustics. Indeed, it is challenging to generate pressures that are sufficiently large to measure the coefficients of the cubic and higher terms in the equation of state. To illustrate this fact, a harmonic signal at $260 \mathrm{~dB} / / 1 \mu \mathrm{~Pa}$ in water corresponds to $\left|\rho / \rho_{0}-1\right|=0.0065$.

A corollary of halting the equation of state at quadratic terms is that any constitutive parameters that are derived from it may be truncated at a consistent level. To formulate Riemann's solution, we need the density as a function of pressure. A Taylor series expansion of the quadratic equation gives

$$
\begin{equation*}
\frac{\rho}{\rho_{0}}=1+\frac{p}{\rho_{0} c^{2}}-\frac{1}{2} \frac{B}{A}\left(\frac{p}{\rho_{0} c^{2}}\right)^{2}+\cdots \tag{13.1.27}
\end{equation*}
$$

Differentiation of Eq. (13.1.26) gives

$$
\begin{equation*}
\frac{\tilde{c}}{c}=\left[1+\frac{B}{A}\left(\frac{\rho}{\rho_{0}}-1\right)\right]^{1 / 2}=1+\frac{B}{2 A}\left(\frac{p}{\rho_{0} c^{2}}\right)+\cdots \tag{13.1.28}
\end{equation*}
$$

In turn, the differential equation relating $v$ to $p$ becomes

$$
\begin{equation*}
\frac{d v}{d p}=\frac{1}{\rho \tilde{c}}=\frac{1}{\rho_{0} c\left(1+\frac{p}{\rho_{0} c^{2}}+\cdots\right)\left[1+\frac{B}{2 A}\left(\frac{p}{\rho_{0} c^{2}}\right)+\cdots\right]} \tag{13.1.29}
\end{equation*}
$$

A consistent series truncation of a binomial series simplifies the differential equation to

$$
\begin{equation*}
\frac{d v}{d p}=\frac{1}{\rho_{0} c}\left[1-\left(\frac{B}{2 A}+1\right)\left(\frac{p}{\rho_{0} c^{2}}\right)+\cdots\right] \tag{13.1.30}
\end{equation*}
$$

Separating variables with the lower limits of the integral set to give $v=0$ if $p=0$ yields

$$
\begin{equation*}
\frac{v}{c}=\left(\frac{p}{\rho_{0} c^{2}}\right)-\frac{1}{2}\left(\frac{B}{2 A}+1\right)\left(\frac{p}{\rho_{0} c^{2}}\right)^{2}+\cdots \tag{13.1.31}
\end{equation*}
$$

[^68]The inverse relation is

$$
\begin{equation*}
\frac{p}{\rho_{0} c^{2}}=\frac{v}{c}+\frac{1}{2}\left(\frac{B}{2 A}+1\right)\left(\frac{v}{c}\right)^{2}+\cdots \tag{13.1.32}
\end{equation*}
$$

The expressions for all quantities derived from the equation of state for a liquid have been truncated on the basis that $|p| /\left(\rho_{0} c^{2}\right)$ is very small. Although the bulk modulus of any gas is much less than that of water, it nevertheless is large, so $|p| /\left(\rho_{0} c^{2}\right)$ for a signal in an ideal gas typically is much less than one. Series expansions of the basic quantities based on taking $|p| /\left(\rho_{0} c^{2}\right)$ to be much less than one are said to constitute small-signal approximations. Waves that have this feature are said to be weakly nonlinear. Acoustical studies seldom need to consider strong nonlinearity.

In such circumstances, the relations for ideal gases and liquid may be combined. To see how this is achieved, consider a Taylor series in powers of $p /\left(\rho_{0} c^{2}\right)$ of the equation of state for an ideal gas, Eq. (13.1.21). When this series is truncated at the first term beyond the expression for linear theory, the result is

$$
\begin{equation*}
\frac{\rho}{\rho_{0}}=1+\frac{p}{\rho_{0} c^{2}}-\frac{\gamma-1}{2}\left(\frac{p}{\rho_{0} c^{2}}\right)^{2}+\cdots \tag{13.1.33}
\end{equation*}
$$

This relation and Eq. (13.1.27) are power series. Changing $B / A$ in the latter to $\gamma-1$ yields the expression for a gas. A description that covers both media is obtained by coefficient of nonlinearity defining a $\beta$ such that

$$
\beta=\left\{\begin{array}{l}
\frac{1}{2}(\gamma+1): \text { Ideal gas }  \tag{13.1.34}\\
\frac{B}{2 A}+1: \text { Liquid }
\end{array}\right.
$$

Because the equations of state now match, the expressions for $\tilde{c}$ and $v$ in terms of $p$ derived from them also will match. The unified constitutive relations that result are

$$
\begin{gather*}
\frac{\rho}{\rho_{0}}=1+\frac{p}{\rho_{0} c^{2}}-(\beta-1)\left(\frac{p}{\rho_{0} c^{2}}\right)^{2}+\cdots \\
\frac{\tilde{c}}{c}=1+(\beta-1)\left(\frac{p}{\rho_{0} c^{2}}\right)+\cdots  \tag{13.1.35}\\
\frac{p}{\rho_{0} c^{2}}=\frac{v}{c}+\cdots
\end{gather*}
$$

The coefficient of nonlinearity is defined as it is in Eq. (13.1.34) because the difference between the nonlinear and linear phase speeds is proportional to $\beta$, specifically

$$
\begin{equation*}
\tilde{c}+v=c+\beta v=c+\beta \frac{p}{\rho_{0} c} \tag{13.1.36}
\end{equation*}
$$

Smallness of $p /\left(\rho_{0} c^{2}\right)$ has the further implication that it usually is acceptable to use $v=p /\left(\rho_{0} c\right)$, which is the same as the relation according to linear theory. In that case, we may shorten the analysis slightly. Instead of solving Eq. (13.1.18) for $v$ and then evaluating the corresponding $p$, we may convert that relation to one that governs $p$ directly,

$$
\begin{equation*}
p=\rho_{0} c \dot{X}(\tau), \quad v=\dot{X}(\tau), \quad \tau=t-\frac{x}{c\left(1+\beta p /\left(\rho_{0} c^{2}\right)\right)} \tag{13.1.37}
\end{equation*}
$$

where the boundary condition is taken as $v=\dot{X}(t)$ at $x=0$.
In summary, four fundamental aspects collectively constitute the small-signal approximation:

- Continuity of particle velocity at a vibrating surface is satisfied at the static reference position of that boundary.
- Liquids and gases are governed by the same equations.
- The pressure and particle velocity for a plane wave propagating in the direction of increasing $x$ are related by $p=\rho_{0} c v$.
- The phase speed along a characteristic is $c+\beta p /\left(\rho_{0} c\right)$.

Each of these simplifies some aspect of an analysis, but none is essential.

### 13.1.5 Quantitative Evaluations

Now that we know $\tilde{c}$ and $v$ as functions of $p$, the next task it to extract profiles and waveforms from the Riemann solution. We will address the boundary value solution in Eq. (13.1.13), but the developments are readily modified to treat the initial value solution in Eq. (13.1.19). It is convenient to nondimensionalize the velocity excitation at $x=0$ by setting

$$
\begin{equation*}
\dot{X}(\tau)=\varepsilon c V(t) \tag{13.1.38}
\end{equation*}
$$

The parameter $\varepsilon$ is the acoustic Mach number we encountered earlier. If we define $V(t)$ such that $\max (|V|)=1$, then $\varepsilon$ is the ratio of the maximum particle velocity to the linear speed of sound. In the small-signal approximation, $\varepsilon$ is very small.

## Graphical Construction

A useful approximation in some situations is a representation of $V(t)$ as a piecewise linear function, that is, a sequence of straight lines with a few locations at which the slope changes. As the signal propagates, waveforms and spatial profiles of pressure retain that attribute if the small-signal approximations apply. Consequently, it is only
necessary to follow the propagation of the phases at which the slope of the initial waveform changes. After those phases are located, the intervals between them remain straight lines.

Suppose $\tau$ is an instant at which the slope of $V(\tau)$ changes. The particle velocity for this phase is $\varepsilon c V(\tau)$, and the small-signal approximation of the pressure is $p(\tau)=\varepsilon \rho_{0} c^{2} V(\tau)$. To construct a waveform, the value of $x$ is fixed. According to Eq. (13.1.36), the phase speed is $c[1+\beta \varepsilon V(\tau)]$. Hence, the time required for this signal to propagate from the origin to the specified $x$ is $x /[c+\beta \varepsilon c V(\tau)]$. By definition, this signal was generated at $x=0$ when $t=\tau$, so the corresponding arrival time of $p(\tau)$ is $\tau+x /[c+\beta \varepsilon c V(\tau)]$. A plot of the waveform observed at $x$ would show this time as the abscissa and $p(\tau)$ as the ordinate. This process is repeated for each value of $\tau$ at which $V(\tau)$ changes slope, after which the waveform is obtained by joining the plotted points with straight lines. The procedure is illustrated in Fig. 13.5 for a triangular pulse whose duration is $T$.


Fig. 13.5 Graphical construction of the waveforms at two spatial locations $x_{A}<x_{B}$. The phase speed of the signal at each instant of slope discontinuity is constant

There are three instants along which the slope changes discontinuously, $\tau_{1}=0$, $\tau_{2}=0.75 T$, and $\tau_{3}=T$. The particle velocity at $\tau_{2}$ is the maximum value $\varepsilon c$. The first plot is the waveform at $x=0$, which is obtained by setting $p=\rho_{0} c \varepsilon V(t)$. Because $v\left(\tau_{1}\right)=v\left(\tau_{3}\right)=0$, the phase speed of these signals is the linear speed of sound, $c$. The time required for $p\left(\tau_{1}\right)$ and $p\left(\tau_{3}\right)$ to propagate to a position $x_{A}>0$ is $x_{A} / c$. The pressure along the characteristic line associated with $\tau_{2}$ is $p=\rho_{0} c^{2} \varepsilon$. The phase speed along this characteristic is constant at $c+\beta \varepsilon c V(\tau)$. The propagation time
required for $p\left(\tau_{2}\right)$ to arrive at $x_{A}$ is $x_{A} /(c+\beta \varepsilon c V(\tau))$, which is less than the travel time for $p\left(\tau_{1}\right)$ and $p\left(\tau_{3}\right)$. Thus, the waveform seems to begin to lean in the sense of earlier time. (If the waveform has an interval in which $p$ is negative, that portion of the waveform will seem to lean in the sense of later time.)

To obtain the waveform at a farther location $x_{B}$ distance, we can evaluate the arrival time for each phase by adding the departure time $\tau$ to $x_{B}$ divided by the phase speed for that pressure, as is done in Fig. 13.5. An alternative is to use the waveform at $x_{A}$ as the reference. The propagation time from $x_{A}$ to $x_{B}$ is $x_{B}-x_{A}$ divided by the phase speed corresponding to the pressure on that characteristic.

At a certain $x$, the arrival time of $p\left(\tau_{2}\right)$ will equal that of $p\left(\tau_{1}\right)$. Such a condition marks the formation of a shock. Equating the two arrival times gives $\tau_{2}+x /(c+\beta \varepsilon c)=\tau_{1}+x / c$. This value of $x$ is the shock formation distance,

$$
\begin{equation*}
x_{\text {shock }}=\frac{c+\beta \varepsilon c}{\beta \varepsilon c} c\left(\tau_{2}-\tau_{1}\right) \tag{13.1.39}
\end{equation*}
$$

(This expression could be simplified by approximating the numerator as $c$ based on $\varepsilon \ll 1$.) The value of $x_{\text {shock }}$ depends on the magnitude and shape of the initial waveform. For now, we regard this as the limiting range for our evaluations.

Evaluation of a spatial profile follows a similar procedure. We select $\tau$ as the phase of a significant feature we wish to follow. The phase speed of this feature is $c+\beta \varepsilon c$. It departed the boundary at time $\tau$. At a designated instant $t$, the propagation time is $t-\tau$, so the feature will be situated at $x=(c+\beta \varepsilon c) t$.

Figure 13.6 depicts this construction for the initial waveform in Fig. 13.5. The first profile describes $t_{A}=\tau_{3}$. This is the instant when $p\left(\tau_{3}\right)$ leaves the boundary, so $x_{3}=0$. Because $p\left(\tau_{1}\right)=0$, its phase speed is $c$ and its location is $x_{1}=c t_{A}$. The propogation time for $p\left(\tau_{2}\right)$ is $t_{A}-\tau_{2}$. According to the Riemann solution, its location

Fig. 13.6 Construction of spatial profiles at various instants

is $x_{2}=c\left(t_{A}-\tau_{2}\right)+\beta \varepsilon c\left(t_{A}-\tau_{2}\right)$. However, the pulse is very brief. Consequently, $c\left(t_{A}-\tau_{2}\right)$ is small, and $\beta \varepsilon c\left(t_{A}-\tau_{2}\right)$ is a negligible addition. Thus, at this early instant the spatial profile differs little from the prediction of linear theory. The second profile is for a much later instant $t_{B}$. The phase speed of $p\left(\tau_{1}=0\right)$ and $p\left(\tau_{3}\right)$ is $c$, so the spatial extent of the pulse remains at $c \tau_{3}$. However, the profile changes shape because the phase speed of $p\left(\tau_{2}\right)$ is greater than $c$. Because $p\left(\tau_{2}\right)$ propagates farther than the beginning and end of the pulse, the profile seems to lean in the sense of increasing $x$. (Regions where the pressure is negative would seem to lean in the sense of decreasing $x$.)

With increasing time, the profile propagates farther, and the $p\left(\tau_{2}\right)$ signal increasingly advances relative to the $p\left(\tau_{1}\right)$ and $p\left(\tau_{3}\right)$ signals. Eventually, it catches up to $p\left(\tau_{1}\right)$, which marks the advent of a shock. The instant $t_{\text {shock }}$ when this condition occurs may be found by equating the locations where $p\left(\tau_{1}\right)$ and $p\left(\tau_{2}\right)$ occur. The shock formation distance $x_{\text {shock }}=c t_{\text {shock }}$ is the same as Eq.(13.1.39).

> EXAMPLE 13.1 The waveforms described in the sketch are alternative velocity inputs at $x=0$. Their only difference is phase-reversal. For each case, determine the shock formation distance, $x_{\text {shock }}$. Then, evaluate each waveform at $x=x_{\text {shock }} / 2$ and $x=x_{\text {shock }}$. The value of $v_{0}$ is sufficiently small to permit considering the propagation to be weakly nonlinear.


Figure 1.

## Significance

Graphical constructions enhance intuitive understanding, which is further enhanced by examining the similarities and differences between the propagation properties of the alternative waveforms.

## Solution

The input functions are the velocity at $x=0$, where $\tau=t$. There are four instants at which the slope of each input changes. For case $\mathrm{A}, v_{1}^{(A)}=0$ at $\tau=0, v_{2}^{(A)}=v_{0}$ at $\tau=T / 4, v_{3}^{(A)}=-v_{0}$ at $\tau=3 T / 4$, and $v_{4}^{(A)}=0$ at $\tau=T$. For case B, the velocities
at $\tau_{2}$ and $\tau_{3}$ are the negative of the value for case A . According to the weakly nonlinear approximations, $p /\left(\rho_{0} c^{2}\right)=v / c$, so the equations of the relevant characteristics are

$$
\begin{array}{cc}
\text { Case A } & \text { Case B } \\
t_{1}=\frac{x}{c} & t_{1}=\frac{x}{c} \\
t_{2}=\frac{T}{4}+\frac{x}{c+\beta v_{0}} & t_{2}=\frac{T}{4}+\frac{x}{c-\beta v_{0}} \\
t_{3}=\frac{3 T}{4}+\frac{x}{c-\beta v_{0}} & t_{3}=\frac{3 T}{4}+\frac{x}{c+\beta v_{0}} \\
t_{4}=T+\frac{x}{c} & t_{4}=T+\frac{x}{c}
\end{array}
$$

A shock forms at the location where characteristics for different values of $\tau$ intersect. The denominator of the $x$ term for $t_{2}$ in case A is bigger than the denominator for $t_{1}$, so there is a value of $x$ at which $t_{2}$ and $t_{1}$ are equal. Similar reasoning suggests that it is possible that $t_{4}=t_{3}$ in case A , whereas in case B the only possibility is that $t_{3}=t_{2}$. These conditions occur at

$$
\begin{aligned}
\text { Case A: } t_{2} & =t_{1} \Longrightarrow \frac{x}{c}-\frac{x}{c+\beta v_{0}}=\frac{T}{4} \Longrightarrow \beta \frac{v_{0}}{c} x=\left(c+\beta v_{0}\right) \frac{T}{4} \\
\text { Case A: } t_{4} & =t_{3} \Longrightarrow \frac{x}{c-\beta v_{0}}-\frac{x}{c}=\frac{T}{4} \Longrightarrow \beta \frac{v_{0}}{c} x=\left(c-\beta v_{0}\right) \frac{T}{4} \\
\text { Case B: } t_{3}=t_{2} & \Longrightarrow \frac{x}{c-\beta v_{0}}-\frac{x}{c+\beta v_{0}}=\frac{T}{2} \Longrightarrow 2 \beta \frac{v_{0}}{c} x=\frac{c^{2}-\left(\beta v_{0}\right)^{2}}{c} \frac{T}{2}
\end{aligned}
$$

The value of $v_{0} / c$ is stated to be sufficiently small to justify using weak shock approximation, so the expression for $x$ in each case reduces to

$$
x_{\text {shock }}=\frac{c T}{4 \beta\left(v_{0} / c\right)}
$$

To plot the waveforms at $x=x_{\text {shock }} / 2$, we substitute $x_{\text {shock }} / 2$ into the expressions for each $t_{j}$ and then plot each time value against the corresponding pressure $\rho_{0} c v\left(\tau_{j}\right)$. The time values are simplified by using the smallness of $v_{0} / c$ to approximate $1 /\left(c \pm \beta v_{0}\right)$ by the first two terms of a binomial series. For example,

$$
\text { At } x=\frac{x_{\text {shock }}}{2}, t_{2} \approx \frac{T}{4}+\frac{\left(1-\beta \frac{v_{0}}{c}\right)}{c} \frac{c T}{8 \beta\left(v_{0} / c\right)}=\frac{T}{8}+\frac{x_{\text {shock }}}{2 c}
$$

The full set of values is

$$
\begin{array}{cc}
\text { Case A }\left(x=x_{\text {shock }} / 2\right) & \text { Case B }\left(x=x_{\text {shock }} / 2\right) \\
t_{1}=\frac{x_{\text {shock }}}{2 c} & t_{1}=\frac{x_{\text {shock }}}{2 c} \\
t_{2}=\frac{T}{8}+\frac{x_{\text {shock }}}{2 c} & t_{2}=\frac{3 T}{8}+\frac{x_{\text {shock }}}{2 c} \\
t_{3}=\frac{7 T}{8}+\frac{x_{\text {shock }}}{2 c} & t_{3}=\frac{5 T}{8}+\frac{x_{\text {shock }}}{2 c} \\
t_{4}=T+\frac{x_{\text {shock }}}{2 c} & t_{4}=T+\frac{x_{\text {shock }}}{2 c}
\end{array}
$$

The waveform corresponding to these instants is shown in Fig. 2.


Figure 2.

The same procedure with $x$ set to $x_{\text {shock }}$ leads to

$$
\begin{array}{cc}
\text { Case A }\left(x=x_{\text {shock }}\right) & \text { Case B }\left(x=x_{\text {shock }}\right) \\
t_{1}=\frac{x_{\text {shock }}}{c} & t_{1}=\frac{x_{\text {shock }}}{c} \\
t_{2}=\frac{x_{\text {shock }}}{c} & t_{2}=\frac{T}{2}+\frac{x_{\text {shock }}}{c} \\
t_{3}=T+\frac{x_{\text {shock }}}{c} & t_{3}=\frac{T}{2}+\frac{x_{\text {shock }}}{c} \\
t_{4}=T+\frac{x_{\text {shock }}}{c} & t_{4}=T+\frac{x_{\text {shock }}}{c}
\end{array}
$$

Figure 3 shows that both waveforms have indeed developed a shock at $x_{\text {shock }}$.


Figure 3.

Although it might seem as though there is a fundamental difference between the signals generated by the alternative inputs, they actually are quite similar. Rather than considering case B to be the negative of case $A$ at $x=0$, the similarities become evident if we compare the behavior of the positive and negative phases. In comparison with case A, the positive phase in case $B$ is retarded in time by $T / 2$, and the negative phase is advanced by $T / 2$. From this perspective, the corresponding phases in each case are altered in the same way as the signal propagates. Both waveforms have the shape of teeth of a sawblade. We will see that this is a tendency that is shared by all oscillatory waveforms.
Computation of Characteristics Graphical construction of waveforms and profiles is impractical if the input is not piecewise linear. It also is inappropriate if we believe that the acoustic Mach number is too large to consider the signal to be weakly nonlinear. The method developed here is a computational version of the graphical algorithm. It allows for evaluation of as many points as necessary to resolve the input waveform, so it can be used in any situation.

We begin with the weakly nonlinear signal. In the graphical method, a value of the phase variable $\tau$ is selected. The initial velocity $v(\tau)$ is used to evaluate the associated pressure $p(\tau)$, from which either $x$ or $t$ is determined by solving the equation for the nonlinear phase speed of that signal. The concept is to use a computer algorithm to perform these operations for many values of $\tau$ in situations where the input function is complicated.

The first step is to discretize the interval during which the input function is nonzero. At $x=0, \tau$ and $t$ are equal, so we let $\tau_{n}$ denote the sampled instants. (Usually we select a uniform increment, so that $\tau_{n}=n \Delta$, but doing so is not essential.) The corresponding set of particle velocities is $v_{n}=\varepsilon c V\left(\tau_{n}\right)$, and the pressures are $p_{n}=\varepsilon \rho_{0} c^{2} V\left(\tau_{n}\right)$.

If we wish to construct the waveform at some point $x_{A}$, we solve Eq.(13.1.36) for the time $t_{n}$ corresponding to the selected values of $x_{A}$ and $\tau_{n}$, which gives

$$
\begin{equation*}
t_{n}=\tau_{n}+\frac{x_{A}}{c\left(1+\beta \varepsilon V\left(\tau_{n}\right)\right)} \tag{13.1.40}
\end{equation*}
$$

Plotting $p_{n}$ as the ordinate and $t_{n}$ as the abscissa for all $n$ yields the waveform. To obtain the profile at a specific instant $t_{A}$, we solve the characteristic's equation for the location $x_{n}$,

$$
\begin{equation*}
x_{n}=\left(1+\beta \varepsilon V\left(\tau_{n}\right)\right) c\left(t_{A}-\tau_{n}\right) \tag{13.1.41}
\end{equation*}
$$

The desired profile results from plotting $p_{n}$ versus $x_{n}$.
Unlike the graphical method, the procedure is readily modified to handle a situation where $\varepsilon \equiv \max (v) / c$ is sufficiently large that the weakly nonlinear approximations might not be valid. It is seldom necessary to consider this possibility for a liquid, so Eq. (13.1.25) for an ideal gas would be the governing equation of state. Rather than using $p_{n}=\rho_{c} c v_{n}$ to evaluate the pressure associated with characteristic $\tau_{n}$, we would use Eq. (13.1.25). The corresponding value $\tilde{c}_{n}$ would be found from Eq. (13.1.22). Characteristic lines in this case are described by the original expres-
sion, Eq.(13.1.13), so the relation between $x$ and $t$ for the signal that departed the boundary at time $\tau$ is

$$
\begin{equation*}
x=\left(t-\tau_{n}\right)\left(\tilde{c}_{n}+v_{n}\right) \tag{13.1.42}
\end{equation*}
$$

In all other respects, the evaluation would proceed as it does when the signal is weakly nonlinear.

EXAMPLE 13.2 A plane wave in air is induced by a harmonically varying velocity at $x=0$. The motion begins at $t=0$, and the velocity amplitude is $0.002 c$. Evaluate the spatial distribution of the pressure signal at $t=50 \mathrm{~T}$, where $T$ is the period of the boundary velocity. Also determine the waveform at $x=50 \lambda$, where $\lambda$ is the wavelength according to linear theory.

## Significance

Application of the general procedure for tracing signals along characteristic lines serves to enhance our ability to implement the method for other problems. Furthermore, harmonic signals are the fundamental building blocks of linear acoustics, so using such a signal as the exemplar will bring to the fore some important phenomena.

## Solution

The excitation is $v=\varepsilon c \sin (\omega \tau) h(t)$ at $x=0$ with $\varepsilon=0.002$ as specified. For reference, $\varepsilon=2\left(10^{-3}\right)$ corresponds to $140 \mathrm{~dB} / / 20 \mu \mathrm{~Pa}$. The fluid is specified to be air, so $\gamma=1.4$. The sound pressure level is quite high. For this reason, we should assess whether it is appropriate to use the small-signal approximations. To do so, let us consider the signal at the instant when the velocity is $\varepsilon c$. According to Eq. (13.1.25), $v / c=0.002$ corresponds to $p=0.0020024 \rho_{0} c^{2}$, whereas the small-signal approximations give $p=0.002 \rho_{0} c^{2}$. The error for the latter value is $0.1 \%$, so application of the small-signal approximations is justified.

There is no need to specify the frequency if we work nondimensionally. To that end, we write the equation for the characteristic as

$$
\omega \tau_{n}=\omega t-\frac{k x}{\left(1+\beta v_{n} / c\right)}
$$

In other words, we take the characteristic value to be $\omega \tau_{n}$ and consider the coordinates in space-time to be $k x$ and $\omega t$.

Before we can select the $\tau_{n}$ values, we must decide the interval to be discretized. The velocity input at $x=0$ begins at $t=0$, at which instant $v=0$. This corresponds to $p=0$, so the characteristic line through the origin is $t-x / c=0$. Thus, at any instant $t$, the signal will extend in space from $x=0$ to $x=c t$. Standard terminology is to say that $x=c t$ is the leading edge. We seek the profile for $t=50(2 \pi / \omega)$, so the range to be sampled is $0 \leq \omega \tau \leq 100 \pi$. To assure that the sampling leads to smooth curves, the sampling rate is selected to be 20 per period, so $\omega \tau_{n}=n(2 \pi / / 20)$,
$n=0,1, \ldots 1000$. In regard to evaluation of waveforms, the acoustic wavelength is $\lambda=2 \pi / k$, so we must determine the waveform at $k x=100 \pi$. It will be proven later that a periodic excitation at the boundary leads to waveforms having the same period. Consequently, we only need to see one period of the waveform. The sampling of $\omega \tau_{n}$ for the evaluation of profiles is adequate for this purpose.

The algorithm is implemented by defining a column vector of $\tau_{n}$ values. The column vector of velocity values at the sampled instants is $v_{n} / c=\varepsilon \sin \left(\omega \tau_{n}\right)$. The pressure corresponding to each $v_{n}$ is $p_{n} /\left(\rho_{0} c^{2}\right)=v_{n} / c$. To find a spatial profile at time $\omega t$, we solve the characteristic definition for the $k x_{n}$ value associated with $p_{n}$ at this $\omega t$, so that $k x_{n}=\left(1+\beta \varepsilon v_{n} / c\right)\left(\omega t-\omega \tau_{n}\right)$. Waveforms result from solving Eq. (1) for $\omega t_{n}$ at location $k x_{n}$ corresponding to pressure $p_{n}$, which gives $\omega t_{n}=$ $\omega \tau_{n}+k x /\left(1+\beta v_{n} / c\right)$.

Figure 1 describes the profile at $\omega t=100 \pi$. (Division of $k x_{n}$ by $2 \pi$ gives distance as a fraction of a linear wavelength. Division of $\omega t_{n}$ by $2 \pi$ gives time as a fraction of the period.) The first plot displays the full extent of the signal. The compressed nature of the $x$ scale makes it difficult to see specific features, so the second graph shows the profile from the boundary outward to five wavelengths. Nothing remarkable is evident-the distribution seems to be the sine curve predicted by linear theory. The third plot shows the region five wavelengths back from the wavefront, $x=c t$. The wave in this region seems to be leaning forward in space. This is a consequence of the positive pressure phase propagating faster than $c$, while the negative phase propagates slower than $c$.


Figure 1.

What would happen if we were to perform the same construction for a much larger time? The leaning effect will be so exaggerated that the computation would yield a multivalued profile. This behavior is exhibited in Fig. 2, which describes the forward part of the spatial profile at $\omega t=300 \pi$. Only one


Figure 2.
pressure can exist at a specific location and instant, so it is evident that there is something wrong. What is missing is the formation of a shock, at which the pressure changes discontinuously. The instant at which the shock forms is that at which a multivalued condition first occurs. Because the positive lobe at the leading edge, $x=c t$, leans forward the most, occurrence of a shock is first manifested at the leading edge as infinite $\partial p / \partial x$. We will investigate shocks in a later section, where we will find that an infinite gradient is a general criterion marking shock formation.

The time required for a positive pressure to propagate to a specific $x$ is less than it is for a negative pressure. Therefore, the leaning appearance of waveforms is in the opposite sense of a spatial profile. This feature is evident in Fig. 3. With increasing $x$, the leaning tendency increases, until a shock forms at a certain distance. The first occurrence of a shock in a waveform is manifested by a vertical tangent at some instant, at which $\partial p / \partial t$ is infinite.


Figure 3.

The waveform has other interesting properties. It is zero until $t=x / c$ and then, it is periodic at the period $T$ of the boundary excitation. Periodicity follows from the fact that the characteristic lines in $x, t$ space corresponding to a periodic $v(\tau)$ form a repeated pattern. This attribute will be proven to be general. A periodic waveform generated by a periodic excitation may be represented as a Fourier series. Thus, another effect of nonlinearity is to create harmonics whose amplitude is dependent on the propagation distance. It might happen that the some harmonics are depleted, while others are generated or enhanced. In any event, propagation leads to an ever increasing level of distortion. Distortion is another effect that will be explored in greater detail.

## Solution of a Nonlinear Algebraic Equation

The method of constructing profiles and waveforms by computing $x$ or $t$ along a characteristic is reliable and readily programmed, so why seek another method? One reason is that analyses that include dissipation or nonplanar effects might be posed in a different form. A more common reason lies in the fact that the interval between the instants at which a waveform is computed is not uniform, that is, a constant increment
$\tau_{n+1}-\tau_{n}$ does not lead to a constant increment $t_{n+1}-t_{n}$. The consequence of the waveform being known discretely at instants that are not uniformly spaced is that standard FFT routines cannot be used to decompose a waveform. This is a serious shortcoming, because the harmonic content of the distorted wave often is of primary interest.

The concept underlying the method we shall now develop is that the Riemann solution may be assembled as a nonlinear equation for $p$ as a function of $x$ and $t$. Such an equation may be solved with the aid of numerical methods. We will derive the method for the boundary value problem, but the procedure may be modified with little effort to address the initial value problem. As before, the boundary condition is $v=\varepsilon c V(t)$ at $x=0$.

Let us begin by considering the procedure when $\varepsilon$ is sufficiently small to warrant considering the signal to be weakly nonlinear. Correspondingly, we set $p=\rho_{0} c^{2} \varepsilon V(t)$ at $x=0$. The same approximation is used to replace $v / c$ in the equation for the characteristics. Thus, the nonlinear plane wave matching the boundary velocity is defined by

$$
\begin{equation*}
p=\varepsilon \rho_{0} c^{2} V\left(t-\frac{x / c}{1+\beta p /\left(\rho_{0} c^{2}\right)}\right) \tag{13.1.43}
\end{equation*}
$$

It is evident that Eq.(13.1.43) is a function relating $x, t$, and $p$. Cases where $p$ and either $x$ or $t$ are specified are equivalent to the previous method, which used constancy of the phase speed along a characteristic. Our interest here is in finding the value of $p$ at specified $x$ and $t$. The equation cannot be solved analytically, so we use numerical methods. It is useful to nondimensionalize the pressure relative to the bulk modulus, such that

$$
\begin{equation*}
q=\frac{p}{\rho_{0} c^{2}} \tag{13.1.44}
\end{equation*}
$$

Correspondingly, the equation to be solved is

$$
\begin{equation*}
G(q, x, t) \equiv q-\varepsilon V\left(t-\frac{x / c}{1+\beta q}\right)=0 \tag{13.1.45}
\end{equation*}
$$

Many methods have been developed to find the roots of a nonlinear algebraic equation. One that works well for the function described above is Newton's method, which is sometimes referred to as the method of tangents. Let $q^{(j)}$ denote the $j$ th iteration. The next iteration is described by

$$
\begin{equation*}
q^{(j+1)}=q^{(j)}-\left.\left(\frac{G(q, x, t)}{\frac{\partial}{\partial q} G(q, x, t)}\right)\right|_{q=q^{(j)}} \tag{13.1.46}
\end{equation*}
$$

The derivative of the $G$ function in Eq. (13.1.45) is

$$
\begin{equation*}
\frac{\partial}{\partial q} G(q, x, t)=1-\varepsilon \dot{V}\left(t-\frac{x / c}{1+\beta q}\right) \frac{\beta x / c}{(1+\beta q)^{2}} \tag{13.1.47}
\end{equation*}
$$

where $\dot{V}$ () denotes the time derivative of the excitation function.
A straightforward modification allows for evaluation of the signal in situations where the signal is strongly nonlinear. In that case, the phase speed as a function of pressure is $\tilde{c}(p)+v(p)$, which are the functions given by Eqs. (13.1.22) and (13.1.24), respectively. The corresponding function representing the Riemann solution is

$$
\begin{equation*}
G(q, x, t) \equiv q-\varepsilon V\left(t-\frac{x}{\tilde{c}\left(\rho_{0} c^{2} q\right)+v\left(\rho_{0} c^{2} q\right)}\right)=0 \tag{13.1.48}
\end{equation*}
$$

Carrying out a numerical solution by Newton's method requires the value of $\partial G / \partial q$ corresponding to the latest iterative value of $q^{(j)}$. The expression for that derivative is complicated by the intricate nature of the expressions for $\tilde{c}(p)$ and $v(p)$. Fortunately, it is adequate to use Eq. (13.1.47) for a weakly nonlinear wave to evaluate $\partial G / \partial q$, because a small error in this quantity merely alters the increment of $q^{(j)}$ in Eq. (13.1.46).

The Newton solver algorithm requires an initial guess, $q^{(0)}$. The fact that $|q| \ll 1$ suggests that a good guess would be the value according to linear theory, which is

$$
\begin{equation*}
q^{(0)}=\varepsilon V\left(t-\frac{x}{c}\right) \tag{13.1.49}
\end{equation*}
$$

This usually will lead to a root in a few iterations, but this assertion only holds if $x$ is not too close to the shock formation distance. A more robust approach is to evaluate $G(q, x, t)$ for a set of values $q_{n}$. Then, the best starting guess is that for which $G\left(q_{n}, x, t\right)$ is closest to zero. A computational operation would implement

$$
\begin{equation*}
\min \left(\left|G\left(q_{n}, x, t\right)\right|\right) \Longrightarrow q^{(0)}=q_{n} \tag{13.1.50}
\end{equation*}
$$

In MATLAB, if $G$ is a vector array of these values, one can find $n$ by writing [ $\left.n, \min \_G\right]=\min (\operatorname{abs}(G))$. The appropriate range of $q_{n}$ values is that which matches the range of $\varepsilon V(t)$ for all possible $t$. For example, if $V(t)=\sin (\omega t)$, then the range of $q_{n}$ would be $-\varepsilon$ to $+\varepsilon$.

To illustrate this procedure, consider the signal in Example 13.2. Let us determine the value of $q$ at $k x=100 \pi$ at the instant $\omega t=101.98 \pi$, which is slightly less than one period after the wavefront arrives at that location. The range of $q_{n}$ values is $-\varepsilon$ to $\varepsilon$, but the range is extended in Fig. 13.7 in order to view the behavior of the $G$ function with greater clarity. The function is sampled at $\Delta q_{n}=\varepsilon / 50$. A scan of the computed values of $G$ shows that it crosses the zero axis between $q_{91}=-4\left(10^{-4}\right)$,
where $G=-2.747\left(10^{-6}\right)$, and $q_{92}=-3.6\left(10^{-4}\right)$, where $G=7.623\left(10^{-6}\right)$. The first $q_{n}$ gives the smaller magnitude of $G$, so the solution algorithm would begin with the initial value $q^{(0)}=-4\left(10^{-4}\right)$. The root is found after one iteration based on a convergence criterion that $\left|q^{(j)}-q^{(j-1)}\right|<\varepsilon / 10^{5}$; it is $q=-3.8944\left(10^{-4}\right)$.

Figure 13.7 also shows a situation where the concept of finding $q$ by solving a nonlinear equation becomes problematic. When $\omega t=k x+1.98 \pi$ at $k x=200 \pi$, which is twice as far from the boundary, there are three values of $q$ for which $G=0$ at the selected $x$ and $t$. This means that three values of $q$ are consistent with the Riemann solution. Hence, the occurrence of multiple roots for $G$ at a selected $x$ and $t$ is another way that shocks are manifested. The Riemann solution is not valid in the presence of shocks. Nevertheless, the method by which shocks are incorporated into an analysis will require the multi-valued portion of the Riemann waveform. Such an evaluation could implement the root search procedure, with the scan initiated near each $q$ value at which the $G$ function changes sign.

Fig. 13.7 Graph of the function $G(q, x, t)$ whose root gives the nondimensional pressure $q=p /\left(\rho_{0} c^{2}\right)$ at $\omega t=k x+1.98 \pi, \varepsilon=0.002$


EXAMPLE 13.3 The particle velocity at $x=0$ is $\varepsilon V(t)=$ $\varepsilon c[\sin (\omega t)+0.5 \sin (2 \omega t)]$ for $t>0$, and it was zero for $t<0$. The fluid is air. Consider the case where $\varepsilon=9\left(10^{-5}\right)$. Compare the pressure waveforms at $k x=800 \pi$ and $k x=1600 \pi$ to the waveform predicted by linear theory. Further, because the excitation is periodic at $T=2 \pi / \omega$, it may be represented as a Fourier series. Construct a graph showing the dependence of the Fourier pressure coefficients on the propagation distance.

## Significance

The evaluation of the signal is unremarkable, but the growth or decay of harmonics is a primary issue for nonlinear acoustics. This evaluation is a useful reminder of spectral techniques that were developed in Chap. 1.

## Solution

We will implement the Newton solver algorithm in nondimensional form because the value of $\omega$ is not given. The value of $\varepsilon$ is sufficiently small to justify considering the wave to be weakly nonlinear. The function to be solved is

$$
G(q, k x, \omega t) \equiv q-\varepsilon\left[\sin \left(\omega t-\frac{k x}{1+\beta q}\right)+0.5 \sin \left(2 \omega t-\frac{2 k x}{1+\beta q}\right)\right] h(t)
$$

The derivative of this function with respect to $q$ is

$$
\frac{\partial}{\partial q} G(q, x, t)=1-\varepsilon\left[\cos \left(\omega t-\frac{k x}{1+\beta q}\right)+\cos \left(2 \omega t-\frac{2 k x}{1+\beta q}\right)\right] \frac{(k x) \beta}{(1+\beta q)^{2}} h(t)
$$

The value of $\beta$ for air is 1.2 .
The value of $q$ predicted by linear theory is the most efficient way to initiate Newton's method. If doing so fails to lead to a convergence, then the search algorithm described by Eq.(13.1.50) may be invoked. (This alternative procedure was not required to obtain the present results.) Thus, the root search is initiated with

$$
q^{(0)}=\varepsilon c[\sin (\omega t-k x)+0.5 \sin (2 \omega t-2 k x)] h(t)
$$

We select an even increment for the time values, $\omega t_{n}=2 \pi n / N, n=0, \ldots, N-1$, in order to exploit FFT technology. The nondimensional fundamental frequency is $\omega \tau=2 \pi$, and the sampling interval is $\omega \Delta t=2 n / N$. Thus, the fundamental frequency of the data is one, and its highest frequency is $N / 2$. Although we will only examine harmonics $m \leq 5$, it is crucial that the highest frequency be greater than the significant frequencies contained in the data. The waveform at $x=0$ contains only two harmonics, so the sampling criterion would be met with a small value of $N$. However, as the signal propagates its steepens in a portion of a period. This raises the Nyquist requirement. Consequently, the evaluations will be carried out with $N=128$.

The procedure followed for each $k x$ is to perform a program loop in which $t_{n}$ is set, and the corresponding value of $q_{n}$ is found by Newton's method. After completion of that loop, the set of $q_{n}$ values is input to an FFT routine. The output of that routine is a set of DFT coefficients that usually must be scaled to obtain Fourier series coefficients $P_{m}$. (In MATLAB, this factor is $2 / N$.) The $P_{m}$ values obtained at each $k x$ are preserved for plotting. A simple check of the computations uses the fact that the mean value of $v(t)$ at $x=0$ is zero. This attribute should be preserved as the signal propagates. (The first value returned by most FFT routines is the mean value, $P_{0}$.)

The waveforms at the requested locations appear in Fig. 1. The abscissa is the retarded time, that is, the elapsed time after the wavefront arrives at $\omega t=k x$. Two periods appear there, but the data for the second period merely periodically replicates that of the first, with $\omega t_{n+N}=\omega t_{n}+2 \pi$ and $p_{n+N}=p_{n}$ for $n=0, \ldots N-1$.


Figure 1.

The waveform at $k x=0$ is the same as the linear solution. Although it has the appearance of being advanced for positive $p$ and retarded for negative $p$, this quality is a consequence of the manner in which the harmonics combine, not the effect of nonlinearity. The distortion of the waveform enhances this initial appearance. At $k x=1600 \pi$, the effect is sufficiently large that the slope of the waveform seems to be nearly vertical at the start of each period.

Waveforms at two values of $k x$ were requested, but such a sample is too sparse for a description of the position dependence of the harmonic amplitudes. Therefore, the waveforms at $k x_{m}=400 \pi, 800 \pi, 1200 \pi$, and $1600 \pi$ were processed. The result is Fig. 2. Only the imaginary parts of the amplitudes are shown because the real parts remain zero. None of the amplitudes are constant. The first and second harmonics are the ones contained in the input, whereas the third and higher harmonics are generated by nonlinearity. There are no dissipation mechanisms, so the total energy in a period is conserved. Typical discussions of the generation of harmonics state that energy is transferred from lower harmonics into higher harmonics as the wave propagates. This graph shows this description to be only partially true. The harmonics that are not directly generated are $n \geq 3$. They do indeed grow as the wave propagates. The directly generated harmonics are $n=1$ and $n=2$, but only the former is depleted, whereas $m=2$ grows slightly.


Figure 2.

The interchange of energy between harmonics and the resulting distortion of the waveform are sensitive to the content of the excitation. To demonstrate this, Fig. 3 shows the waveforms that result from shifting the second harmonic of the input by $180^{\circ}$, so that $v=\varepsilon[\sin (\omega t)-0.5 \sin (2 \omega t)]$ at $x=0$. The interference of the harmonics of $p(t)$ at $x=0$ is such that the starting waveform has less of the leaning appearance than it did in the previous case. The nonlinear advancement
and retardation are the same as in the previous case. However, the overall leaning appearance of the initial waveform is less noticeable than it was in the previous case. The consequence is that the maximum slope of the waveform at $k x=1600 \pi$ is reduced, which suggests that shocks will form at a larger distance.


Figure 3.

The harmonic content of these signals is described in Fig. 4. Reversal of the second harmonic at $x=0$ has a rather drastic effect on the growth and decay of harmonic amplitudes. Increasing propagation distance leads to depletion of the $m=2$ harmonic. The energy lost from that harmonic is transferred into the other harmonics, including the first. The growth of the higher harmonics is similar to the previous case, but now the fourth harmonic grows more rapidly than the third.


Figure 4.

### 13.2 Effects of Nonlinearity

The examples disclosed that some interesting phenomena arise in nonlinear plane waves. Two of particular interest are harmonic distortion and the occurrence of shocks. Both features are encountered in other types of waves. Both are sufficiently important to warrant a more detailed analysis, which is what we shall now do.

### 13.2.1 Harmonic Generation

A periodic excitation at the boundary leads to a periodic waveform. This is so because the characteristics form a periodic pattern parallel to the $t$-axis. Periodicity is readily
proven. Two phases separated by a period are $\tau$ and $\tau+T$. The arrival time of these phases at a specified $x$ is $t$ and $t+\Delta t$, where

$$
\begin{equation*}
t=\tau+\frac{x}{c+\beta \varepsilon V(\tau)}, \quad t+\Delta t=(\tau+T)+\frac{x}{c+\beta \varepsilon V(\tau+T)} \tag{13.2.1}
\end{equation*}
$$

Because $V(\tau)$ is periodic, $V(\tau+T)=V(\tau)$, so it must be that $\Delta t=T$. A corollary of periodicity is that the signal may be expressed as a Fourier series. Rather than using an FFT to process discretized data, as we did in Example 13.3, let us consider a formal analysis of the Fourier coefficients.

It is convenient to measure time relative to the signal's arrival at position $x$. Because the pressure at the wavefront is zero, the arrival time is $x / c$. Thus, the time variable is the retarded time $t^{\prime}$, where

$$
\begin{equation*}
t^{\prime}=t-\frac{x}{c} \tag{13.2.2}
\end{equation*}
$$

We consider the weakly nonlinear wave generated by a periodic boundary vibration $v=\varepsilon c V(t)$ at $x=0$. We further require that $V(t)=0$ for $t<0$, in order that the velocity be continuous at the initial instant. This condition is added because a discontinuity in the waveform means the existence of a shock, which invalidates using the Riemann solution throughout a period.

A periodic function that is zero at $t=0$ may be represented as a Fourier sine series. The period of the excitation is $T$, so the fundamental frequency is $\omega=2 \pi / T$. A Fourier sine series whose independent variable is $t^{\prime}$ is used,

$$
\begin{equation*}
p=\sum_{n=1}^{\infty} P_{n} \sin \left(n \omega t^{\prime}\right), \quad t^{\prime}>0 \tag{13.2.3}
\end{equation*}
$$

Concurrently, the Riemann solution in Eq. (13.1.37) also describes $p$, so it must be that

$$
\begin{equation*}
\sum_{n=1}^{\infty} P_{n} \sin \left(n \omega t^{\prime}\right)=\rho_{0} c^{2} \varepsilon V(\tau), \quad \tau=t^{\prime}+\frac{x}{c} \beta \varepsilon V(\tau) \tag{13.2.4}
\end{equation*}
$$

The relation for $\tau$ differs from the earlier form by application of a binomial series expansion to clear the denominator, based on the fact that $|\varepsilon| \ll 1$.

The Fourier pressure coefficients are found by applying the orthogonality property of sine functions to Eq. (13.2.4). The first period following $t^{\prime}=0$ covers the interval $0<t^{\prime}<T$. Therefore, the coefficients are given by

$$
\begin{equation*}
P_{n}=\rho_{0} c^{2} \varepsilon \frac{2}{T} \int_{0}^{T} V(\tau) \sin \left(n \omega t^{\prime}\right) d t^{\prime} \tag{13.2.5}
\end{equation*}
$$

Because $V(\tau)$ is an arbitrary function, we use Eq. (13.2.4) to change the integration variable from $t^{\prime}$ to $\tau$. Doing so gives

$$
\begin{equation*}
P_{n}=\rho_{0} c^{2} \varepsilon \frac{2}{T} \int_{0}^{T} V(\tau) \sin \left[n \omega\left(\tau-\frac{x}{c} \beta \varepsilon V(\tau)\right)\right]\left[1-\frac{x}{c} \beta \varepsilon \dot{V}(\tau)\right] d \tau \tag{13.2.6}
\end{equation*}
$$

The factor of $V(\tau)$ is a perfect differential, which suggests an integration by parts. Because $V(0)=V(T)=0$, this gives

$$
\begin{equation*}
P_{n}=\rho_{0} c^{2} \varepsilon \frac{2}{n \omega T} \int_{0}^{T} \dot{V}(\tau) \cos \left[n \omega\left(\tau-\frac{x}{c} \beta \varepsilon V(\tau)\right)\right] d \tau \tag{13.2.7}
\end{equation*}
$$

The periodic nature of the integrand allows the integration range to be shifted to $-T / 2<\tau<T / 2$. Furthermore, because $V(\tau)$ is an odd function relative to $\tau=0$, $\dot{V}(\tau)$ is even, so that the product of $\dot{V}(\tau)$ and the cosine term also is even. Therefore, the integration interval may be reduced to $0<\tau<T / 2$ with the result of the integral doubled. The result is

$$
\begin{equation*}
P_{n}=\rho_{0} c^{2} \varepsilon \frac{2}{n \pi} \int_{0}^{T / 2} \dot{V}(\tau) \cos \left[\frac{2 n \pi}{T}\left(\tau-\frac{x}{c} \beta \varepsilon V(\tau)\right)\right] d \tau \tag{13.2.8}
\end{equation*}
$$

It might be desirable to compare these $P_{n}$ values to those obtained from FFT analysis of a waveform, which is the procedure in the previous example. To do so, it is necessary to be cognizant of the fact that the FFT gives the coefficients of a complex Fourier series, whereas Eq. (13.2.8) gives the coefficients of a Fourier sine series.

Equation (13.2.8) may be used to describe the propagation of any initial waveform for which $p=0$ at the beginning of a period. The expression seems to be unpromising for an analytical integration, but it can be evaluated numerically. Harmonic excitation is one case where an analytical result can be obtained. If $V(\tau)=\sin (\omega \tau)$, then an identity for the product of cosines, accompanied by the change of variables $\xi=\omega \tau$, leads to

$$
\begin{align*}
P_{n}= & \rho_{0} c^{2} \varepsilon \frac{2}{n \pi} \int_{0}^{\pi} \cos (\xi) \cos [n(\xi-k x \beta \varepsilon \sin (\xi))] d \xi \\
\equiv & \rho_{0} c^{2} \varepsilon \frac{1}{n \pi} \int_{0}^{\pi}\{\cos [(n-1) \xi-n k x \beta \varepsilon \sin (\xi)] \\
& +\cos [(n+1) \xi-n k x \beta \varepsilon \sin (\xi)]\} d \xi \tag{13.2.9}
\end{align*}
$$

Both parts of the integrand match a fundamental formula for Bessel functions. ${ }^{8}$ From it, we find that

$$
\begin{equation*}
\frac{P_{n}}{\rho_{0} c^{2}}=\frac{\varepsilon}{n}\left[J_{n-1}(n \beta \varepsilon k x)+J_{n+1}(n \beta \varepsilon k x)\right] \equiv \frac{2}{n \beta k x} J_{n}(n \beta \varepsilon k x) \tag{13.2.10}
\end{equation*}
$$

[^69]where the last form stems from a recurrence relation for the Bessel functions. A more meaningful form results when distance is described relative to the shock formation distance of an initially harmonic plane wave, $\left(x_{\text {shock }}\right)_{\text {sine }}$. Shock formation is the subject of the next subsection. It will be found that $\left(x_{\text {shock }}\right)_{\text {sine }}=1 /(\varepsilon \beta k)$, so that
\[

$$
\begin{equation*}
\frac{P_{n}}{\rho_{0} c^{2}}=\frac{2 \varepsilon}{n \sigma} J_{n}(n \sigma) \tag{13.2.11}
\end{equation*}
$$

\]

The symbol $\sigma$ is the distance ratio,

$$
\begin{equation*}
\sigma=\frac{x}{\left(x_{\text {shock }}\right)_{\text {sine }}}=\beta \varepsilon k x \tag{13.2.12}
\end{equation*}
$$

The corresponding series representation is

$$
\begin{equation*}
p=\rho_{0} c^{2} \varepsilon \sum_{n=1}^{\infty} \frac{2}{n \sigma} J_{n}(n \sigma) \sin \left(n \omega t^{\prime}\right) \tag{13.2.13}
\end{equation*}
$$

Because this analysis began with the Riemann solution, the basic limitation to regions where $x<\left(x_{\text {shock }}\right)_{\text {sine }}$, or $\sigma<1$, applies equally to the Fourier series.

The Fourier series analysis was first worked out by Fubini-Ghiron in 1935. ${ }^{9}$ A remarkable aspect is the simplicity of the result, for it tells us that the harmonic amplitudes depend on only two parameters: the harmonic number and distance relative to the position where a shock first forms. However, this property is specific to a sinusoidal excitation function; other excitations will lead to dependence on additional parameters.

Figure 13.8 plots the amplitudes. We see that the fundamental harmonic amplitude lessens with increasing distance, whereas the higher harmonic amplitudes grow. Another important property of a sinusoidal excitation is that the amplitudes at any $x$ decrease as the harmonic number increases. This brings up an ambiguity regarding the description of harmonics. The terminology used here is to refer to $n=1$ as the fundamental or first harmonic, $n=2$ as the second harmonic, and so on. However, one can find many technical works wherein $n=1$ is referred to as the fundamental, $n=2$ is the first harmonic, and $n=3$ is the second harmonic. This usage is consistent with terminology for music, but we shall avoid it.

Simpler expressions for the range dependence of the amplitudes may be obtained if we restrict our attention to the region close to the source at $x=0$. Application to Eq. (13.2.13) of the series approximation of Bessel functions gives

[^70]

Fig. 13.8 Plot of the Fourier series coefficients for the signal generated by a sinusoidal excitation at $x=0$. The distance $x_{\text {shock }}$ is the location where a shock forms, and the analysis is not valid beyond that distance

$$
\begin{gather*}
\frac{P_{1}}{\rho_{0} c^{2} \varepsilon}=1-\frac{1}{8}\left(\frac{x}{x_{\text {shock }}}\right)^{2}+\ldots \\
\frac{P_{2}}{\rho_{0} c^{2} \varepsilon}=\frac{1}{2}\left(\frac{x}{x_{\text {shock }}}\right)+\ldots, \frac{P_{3}}{\rho_{0} c^{2} \varepsilon}=\frac{3}{8}\left(\frac{x}{x_{\text {shock }}}\right)^{2}+\ldots \tag{13.2.14}
\end{gather*}
$$

This confirms quantitatively the trend that is evident in Fig. 13.8 that the primary effect near the source, $x=0$, is growth of the second harmonic linearly with propagation distance. The depletion of the fundamental, as well as the growth of third harmonic, depends quadratically on the propagation distance. All of these effects are weak until $x$ is a substantial fraction of $x_{\text {shock }}$.

In view of the fact that we could have obtained the data in Fig. 13.8 from FFT analysis, as we did in Example 13.3, one might think that Fubini-Ghiron's development is important only from a historical perspective. This is far from the truth. For one, many evaluations would be required before we realized that the harmonic amplitudes depend on only two parameters. In addition, the availability of a formula for the amplitudes allows us to delve further into basic properties. One of particular interest is energy conservation.

Consider a temporally periodic wave in the region between two planes at $x_{1}$ and $x_{2}$, both of which are less than $x_{\text {sh }}$. As a consequence of periodicity of the pressure and particle velocity, the sum of the kinetic energy and potential energy per unit volume at any location $x$ is the same at the end of any period as it was at the beginning. Thus, the total energy in the region between the two planes is conserved over any period. It follows that the average energy transported into this region at $x_{1}$ in an interval of a period must equal the average energy that flows out of this region at $x_{2}$. The cross sectional areas at $x_{1}$ and $x_{2}$ are identical and the locations are arbitrary, so it must be that the time-averaged intensity in the propagation direction is constant.

The time-averaged intensity is the same basic expression as it is in linear theory,

$$
\begin{equation*}
\left(I_{x}\right)_{\mathrm{av}}=\frac{1}{T} \int_{t_{0}-T / 2}^{t_{0}+T / 2} p v d t \tag{13.2.15}
\end{equation*}
$$

where $t_{0}$ is arbitrary. The order of magnitude of $p$ and $v$ is $\varepsilon$, so the intensity is $O\left(\varepsilon^{2}\right)$. However, as we proceed some higher-order terms might arise. To assure that we consistently account for all such terms, we shall employ the small-signal approximations, but retain the quadratic terms in the relation between $v$ and $p$. Equation (13.1.35) describes this relation, so we have

$$
\begin{equation*}
\left(I_{x}\right)_{\mathrm{av}}=\frac{c}{T} \int_{t_{0}-T / 2}^{t_{0}+T / 2} p\left[\frac{p}{\rho_{0} c^{2}}-\frac{1}{2} \beta\left(\frac{p}{\rho_{0} c^{2}}\right)^{2}\right] d t \tag{13.2.16}
\end{equation*}
$$

The series is defined in terms of the delayed time $t^{\prime}$, Correspondingly, the intensity may be computed by integrating over $T / 2<t^{\prime}<3 T / 2$. (This is the first full interval centered on an instant when $p=0$.) This changes the preceding expression to

$$
\begin{equation*}
\left(I_{x}\right)_{\mathrm{av}}=\frac{\rho_{0} c^{3}}{T} \int_{T / 2}^{3 T / 2}\left[\left(\frac{p}{\rho_{0} c^{2}}\right)^{2}-\frac{1}{2} \beta\left(\frac{p}{\rho_{0} c^{2}}\right)^{3}\right] d t^{\prime} \tag{13.2.17}
\end{equation*}
$$

The pressure is an odd function with respect to $t^{\prime}=T$, which means that $p\left(t^{\prime}\right)=$ $-p\left(2 T-t^{\prime}\right)$. Therefore $p\left(t^{\prime}\right)^{3}$ also is an odd function. This means that its contribution to the integral from $T / 2<t^{\prime}<T$ will be the negative of its contribution from $T<t^{\prime}<3 T / 2$. The consequence is that the contribution of $p^{3}$ vanishes. Thus the time-averaged intensity is given by the same expression as it is in linear theory. Periodicity permits adjusting the interval to $0<t^{\prime}<T$, so that

$$
\begin{equation*}
\left(I_{x}\right)_{\mathrm{av}}=\frac{1}{T \rho_{0} c} \int_{0}^{T} p^{2} d t^{\prime} \tag{13.2.18}
\end{equation*}
$$

We wish to prove that the Riemann solution for any excitation leads to a timeaveraged intensity that is independent of $x$. Doing so is much easier if we limit consideration to weakly nonlinear signals. In that case the pressure is related to the boundary velocity by $p(\tau)=\rho_{0} c^{2} \varepsilon V(\tau)$, and $\tau$ is related to $t^{\prime}$ by Eq. (13.2.4). The latter relation is used to convert the integration variable from $t^{\prime}$ to $\tau$, which leads to

$$
\begin{align*}
\left(I_{x}\right)_{\mathrm{av}} & =\frac{1}{T \rho_{0} c} \int_{0}^{T} p^{2} d\left[\tau-\frac{x}{c} \beta \varepsilon V(\tau)\right] \\
& =\frac{\rho_{0} c^{3}}{T} \varepsilon^{2} \int_{0}^{T} V(\tau)^{2}\left[1-\frac{x}{c} \beta \varepsilon \dot{V}(\tau)\right] d \tau \\
& =\frac{\rho_{0} c^{3}}{T} \varepsilon^{2}\left[\int_{0}^{T} V(\tau)^{2} d \tau-\left.\frac{x}{3 c} \beta \varepsilon V(\tau)^{3}\right|_{0} ^{T}\right] \tag{13.2.19}
\end{align*}
$$

Because $V(\tau)$ is periodic, the cubic term vanishes. Furthermore, the integrand depends only on the nature of $V(\tau)$, so it is independent of $x$. This proves that $\left(I_{x}\right)_{\mathrm{av}}$ is independent of $x$, being

$$
\begin{equation*}
\left(I_{x}\right)_{\mathrm{av}}=\frac{\rho_{0} c^{3}}{T} \varepsilon^{2} \int_{0}^{T} V(\tau)^{2} d \tau \tag{13.2.20}
\end{equation*}
$$

An alternative form of this expression results from Parseval's theorem, Eq. (1.4.12), which describes the mean-squared pressure corresponding to a Fourier series. The fact that the integration is performed over a period of the $\tau$ variable, not time $t$, is irrelevant. Thus, the time-averaged intensity is related to the Fourier series coefficients of $V(\tau)$ by

$$
\begin{equation*}
\left(I_{x}\right)_{\mathrm{av}}=\frac{1}{2} \rho_{0} c^{3} \varepsilon^{2} \sum_{n=1}^{\infty}\left|V_{n}\right|^{2} \tag{13.2.21}
\end{equation*}
$$

Because the derivation was restricted to weakly nonlinear waves, this expression actually only is accurate to $O\left(\varepsilon^{2}\right)$. An analysis that does not invoke the weak nonlinearity would confirm that the Riemann solution describes a signal in which power flows constantly to the location where a shock forms, regardless of the magnitude of the signal.

### 13.2.2 Shock Formation

We have seen three manifestations of the occurrence of a shock: intersection of characteristic lines, multivaluedness of the function whose roots give the pressure, and vertical tangency of a waveform or spatial profile. Any of these three features may be used as the basis for a quantitative analysis of the distance where a shock forms. The result from each is the same, but the differences in the analysis demonstrate different techniques for analyzing other systems.

We begin with an excitation at $x=0$ that is a piecewise linear waveform like the one in Example 13.1. Whereas we previously tracked the propagation of the phases at which $p$ is maximum and zero, we now consider signals that depart the boundary at two arbitrary instants $\tau_{1}$ and $\tau_{2}>\tau_{1}$; the corresponding pressures are $p_{1}$ and $p_{2}$. The phase speeds associated with these signals are $\tilde{c}_{1}+v_{1}$ and $\tilde{c}_{2}+v_{2}$. The equations for the two characteristics are

$$
\begin{equation*}
\tau_{1}=t-\frac{x}{\tilde{c}_{1}+v_{1}}, \quad \tau_{2}=t-\frac{x}{\tilde{c}_{2}+v_{2}} \tag{13.2.22}
\end{equation*}
$$

At the intersection of these lines, the values of $x$ and $t$ are the same. Because $\tau_{2}>\tau_{1}$, such an intersection can occur only if the slope $d x / d t=1 /\left(\tilde{c}_{2}+\beta v_{2}\right)$ of the $\tau_{2}$ characteristic is less than the slope $1 /\left(\tilde{c}_{1}+\beta v_{1}\right)$ of the $\tau_{1}$ characteristic. We have seen that both $\tilde{c}$ and $v$ increase monotonically with increasing $p$. It follows that the characteristic lines will intersect only if $p_{2}>p_{1}$. This is another way of saying that a shock forms because a later high-pressure phase catches up to an earlier lower
pressure phase. To solve for $x$, we take the difference of the two equations, which leads to

$$
\begin{equation*}
x=\frac{\left(\tilde{c}_{1}+v_{1}\right)\left(\tilde{c}_{2}+v_{2}\right)\left(\tau_{2}-\tau_{2}\right)}{\left(\tilde{c}_{2}+v_{2}\right)-\left(\tilde{c}_{1}+v_{1}\right)} \tag{13.2.23}
\end{equation*}
$$

The shock formation distance, $x_{\text {shock }}$, in general is the minimum value of $x$ at which a shock occurs at any instant. Rather than solving this relation for a variety of $\tau_{1}, \tau_{2}$ pairs, we can identify directly which pair of characteristics intersect at the smallest $x$. Specifically, the $\tau_{2}$ characteristic should be the one for which $\tilde{c}_{2}+v_{2}$ is the largest value, and the $\tau_{1}$ characteristic should be the one for which $\tilde{c}_{1}+v_{1}$ is the smallest. Equation (13.1.39) is the result for a simple triangle waveform because in that case, $p_{2}$ is the maximum and $p_{1}=0$.

Equation (13.2.23) applies if $p(\tau)$ at $x=0$ is piecewise linear. If it is a continuously differentiable function, it is more useful to consider two adjacent instants by letting $\tau_{1}=\tau$ and $\tau_{2}=\tau+d \tau$. Then, $\left(\tilde{c}_{2}+v_{2}\right)-\left(\tilde{c}_{1}+v_{1}\right)$ becomes $d(\tilde{c}+v)$. Furthermore, $\tilde{c}+v$ is a function of $p$, and $p$ is a function of $x$ and $\tau$. These observations lead to identification of the condition for formation of a shock,

$$
\begin{equation*}
x=\frac{(\tilde{c}+v)^{2}}{\left[\frac{d}{d p}(\tilde{c}+v)\right] \frac{\partial p}{\partial \tau}} \tag{13.2.24}
\end{equation*}
$$

We know that $\tilde{c}+v$ increases monotonically with increasing $p$, so that $d(\tilde{c}+v) /$ $d p>0$. Therefore, we conclude that a shock occurs in the portion of the waveform for which $p$ increases with increasing $t$.

Equation (13.2.24) is valid for a signal at any amplitude (within the limits of the assumptions of adiabatic propagation in an inviscid fluid). In acoustic applications, the acoustic Mach number $\varepsilon$ is small, so Eq.(13.1.36) may be used to describe the phase speed. Doing so gives

$$
\begin{equation*}
x=\frac{c\left(1+\beta \frac{p}{\rho_{0} c^{2}}\right)}{\beta \frac{d}{d \tau}\left(\frac{p}{\rho_{0} c^{2}}\right)} \tag{13.2.25}
\end{equation*}
$$

In the small-signal approximation $p /\left(\rho_{0} c^{2}\right)=v / c=\varepsilon V(\tau)$. Thus, when terms whose order of magnitude is $\varepsilon$ or smaller are dropped, the preceding becomes a simple expression for $x$. This value depends on the value of $\tau$. The shock formation distance is defined to be the smallest $x$ at which characteristics intersect because we must assume that the basic formulation ceases to be valid in a region where a shock exists at any instant. The minimum distance, $x_{\text {shock }}$, corresponds to the positive maximum of $d V / d \tau$. Thus,

$$
\begin{equation*}
x_{\text {shock }}=\frac{c}{\beta \varepsilon \max \left(\frac{d V}{d t}\right)} \tag{13.2.26}
\end{equation*}
$$

This expression confirms that the shock forms at the phase of the waveform that changes most rapidly. It also tells us that the distance at which a shock first forms is inversely proportional to the magnitude of the excitation.

Our second encounter with shock formation occurred when we developed a numerical procedure to evaluate the pressure. This entailed finding the value $q \equiv p /\left(\rho_{0} c^{2}\right)$ that satisfies $G(q, x, t)=0$. Figure 13.7 describes this function at $x<x_{\text {shock }}$, in which case there is one root, and at $x>x_{\text {shock }}$, in which case there are three roots. As $x$ is increased from the first value, the slope of $G v s . q$ at the root decreases. The transition from one root to three is marked by a horizontal tangent at the root. This suggests that the condition leading to $x_{\text {shock }}$ is marked by

$$
\begin{equation*}
\frac{\partial}{\partial q} G(q, x, t)=0 \tag{13.2.27}
\end{equation*}
$$

This derivative is given by Eq. (13.1.47), so we set

$$
\begin{equation*}
1-\left.\varepsilon \frac{d V}{d \tau}\right|_{\tau=t-\frac{x / c}{1+\beta q}}\left[\frac{x}{c} \frac{\beta}{(1+\beta q)^{2}}\right]=0 \tag{13.2.28}
\end{equation*}
$$

Because the order of magnitude of $q$ is $\varepsilon$, it is acceptable to approximate the denominator terms as one. The smallest value of $x$ at which this condition occurs corresponds to the largest value of $d V / d \tau$, regardless of when it occurs. The resulting expression for $x_{\text {shock }}$ is the same as Eq. (13.2.26).

Another way in which we may locate a shock is by explicitly searching for a vertical tangency of the waveform or spatial profile. This property is recognizable by the fact that the interval in space for a profile, or in time for a waveform, between the characteristics associated with $\tau$ and $\tau+d \tau$ decreases as they come closer to intersecting. We shall pursue this analysis for a waveform and leave the corresponding analysis of a profile for Exercise 13.13. The first step is to derive an expression for $\partial p / \partial t$. The substitution $p=\rho_{0} c^{2} q$ somewhat simplifies differentiation of the smallsignal approximation in Eq. (13.1.43), which gives

$$
\begin{equation*}
\left.\frac{\partial q}{\partial t} \equiv \varepsilon \frac{d V}{d \tau}\right|_{\tau=t-\frac{x / c}{1+\beta q}}\left[1+\frac{x / c}{(1+\beta q)^{2}} \beta \frac{\partial q}{\partial t}\right] \tag{13.2.29}
\end{equation*}
$$

from which we find that

$$
\begin{equation*}
\left[\left.\frac{x / c}{(1+\beta q)^{2}} \beta \varepsilon \frac{d V}{d \tau}\right|_{\tau=t-\frac{x / c}{1+\beta q}}-1\right] \frac{\partial q}{\partial \tau}=-\left.\varepsilon \frac{d V}{d \tau}\right|_{\tau=t-\frac{x / c}{1+\beta q}} \tag{13.2.30}
\end{equation*}
$$

For a vertical tangency, we seek the condition that leads to $\partial q / \partial t$ being infinite. This occurs if the bracketed term is zero. The minimum $x$ at which this condition occurs corresponds to the maximum value of $d V / d t$, so a vertical tangent first occurs at

$$
\begin{equation*}
\frac{x}{c}=\frac{(1+\beta q)^{2}}{\beta \varepsilon \max \left(\frac{d V}{d \tau}\right)} \tag{13.2.31}
\end{equation*}
$$

Dropping $\beta q$ in comparison with 1 yields Eq.(13.2.26).
Identification of the location where a shock forms is possible by three alternative methods because of the manner in which the Riemann solution is posed. For other systems, these approaches might not be equally viable. Indeed, if the solution is the result of a numerical analysis, the only suitable approach might be a numerical evaluation of the slope of a waveform or profile.

The value of $x_{\text {shock }}$ depends on the nature of the input function $V(\tau)$. The typical concern is a sine function, $V(\tau)=\sin (\omega \tau-\phi)$. In that case, $\max (d V / d \tau)=\omega$, so that

$$
\begin{equation*}
\left(x_{\text {shock }}\right)_{\text {sine }}=\frac{1}{\varepsilon \beta k} \tag{13.2.32}
\end{equation*}
$$

For reference values, a harmonic excitation at 1 kHz in air at $140 \mathrm{~dB} / / 20 \mu \mathrm{~Pa}$ corresponds to $\varepsilon=0.0020$ and $x_{\text {shock }}=23 \mathrm{~m}$, while a $200 \mathrm{~dB} / / 1 \mu \mathrm{~Pa}$ signal at 1 kHz underwater corresponds to $\varepsilon=6.5\left(10^{-5}\right)$ and $x_{\text {shock }}=1.04 \mathrm{~km} .{ }^{10}$ These values may be scaled for other signal amplitudes and frequencies, so that a 1 MHz signal in water at $\varepsilon=6.5\left(10^{-5}\right)$ shocks at 1.04 m , while the same signal at $\varepsilon=6.5\left(10^{-6}\right)$ forms a shock at 10.4 m .

### 13.2.3 Propagation of Weak Shocks

Propagation of a wave beyond the location where a shock first forms causes its spatial profile or waveform to develop an extremely steep slope over a very small interval. This condition is depicted in Fig. 13.9. Within this interval, shear stresses grow, which means the some accounting must be given to the effects of viscosity and heat conduction. Model field equations that account for the role of viscosity and other nonideal effects would be much more difficult to derive and solve. Furthermore, the specific details of the small-scale region within which the signal changes drastically might be of little interest if we can find a simpler way to describe how this phenomenon affects the overall propagation features. The formulation that meets this objective replaces the actual velocity dependence by a model in which the velocity

[^71]

Fig. 13.9 Actual pressure profile and waveform at a shock and the idealized model. The latter replaces extremely steep gradients and high accelerations with discontinuous changes in space and time
changes discontinuously. Figure 13.9 shows that the actual pressure profile changes over a distance that is much less than a wavelength in space. It also shows that the actual waveform changes over an interval that is much shorter than a period. If we do not closely inspect the signal, the model of a discontinuous change is a good approximation. A discontinuity of the pressure in either space or time is an artifice, but one implies the other. This is so because Euler's equation and the equation of state require that $\partial p / \partial t$ is infinite if $\nabla p$ is infinite.

The term "shock" is synonymous with discontinuity. We have already used $x_{\text {shock }}$ to denote the distance at which the wave first develops a vertical tangency, so $x_{\text {dis }}(t)$ will denote the location of the discontinuity in a spatial profile at an arbitrary instant. For a wave propagating in the sense of increasing $x$, it must be that $x_{\text {dis }} \geq x_{\text {shock }}$. Similarly, $t_{\text {dis }}(x)$ is the instant at which a shock is observed in the waveform at an arbitrary location. Other parameters are described in Fig. 13.9. The position infinitesimally behind the discontinuity is denoted as $x_{\text {dis }}^{-}$and the corresponding pressure is $p^{-}$, whereas the position infinitesimally ahead of the discontinuity is $x_{\text {dis }}^{+}$and the corresponding pressure is $p^{+}$. A minor confusion results from the usage of + and - to denote when and where a feature occurs, rather than relative size. Specifically, we have $p^{-}>p^{+}, t^{-}>t^{+}$, and $x^{-}<x^{+}$. In terms of the characteristic variable, because $p^{+}$is ahead of $p^{-}$in space, it must be that $p^{+}$departed from the boundary at an earlier time, so that $\tau^{-}>\tau^{+}$. Multiple discontinuities may exist, but each is treated individually.

## Rankine-Hugoniot Jump Conditions

If the model of a discontinuous pressure change at a shock is to be useful, we must determine how the discontinuity moves in space. This analysis modifies the derivation of the conservation principles to include a discontinuous change of the state variables. The domain we consider is a control volume contained between two fixed locations, $x_{1}<x_{\text {dis }}<x_{2}$. Figure 13.10 shows this control volume at two instants $t$ and $t+d t$. The position of the discontinuity is $x_{\text {dis }}$ at each instant. The field is independent of


Fig. 13.10 A stationary control volume surrounding a moving shock
position transverse to the propagation distance, so the cross section may be considered to be a unit area.

In Fig. 13.10(a), a chunk of fluid is situated to the left of the control volume. This chunk contains the fluid that will enter the control volume during the interval from $t$ to $t+d t$. The width of this chunk is $d s_{\mathrm{in}}$. Because this width is infinitesimal, the velocity of all particles contained in it is essentially $v\left(x_{1}, t\right)$. By definition, the left face of this chunk arrives at $x_{1}$ in an elapsed time of $d t$, so $d s_{\text {in }}=v\left(x_{1}, t\right) d t$. In Fig. 13.10(b), the chunk of fluid to the right of the control consists of the particles that exited in the interval from $t$ to $t+d t$. As a consequence of $d s_{\text {out }}$ being infinitesimal, the velocity of these particles is essentially $v\left(x_{2}, t\right)$. Hence, $d s_{\text {out }}=v\left(x_{2}, t\right) d t$.

The principles we shall consider are conservation of mass, momentum and impulse, and energy and work. The formulation of each balance principle is essentially the same. Let $\mathcal{H}$ denote the quantity that is the subject of the principle. Thus, the mass per unit volume is $\mathcal{H}=\rho$, the momentum per unit volume is $\mathcal{H}=\rho v$, and the energy per unit volume is $\mathcal{H}=(1 / 2) \rho v^{2}+U$, where $U$ is the potential energy per unit volume stored within the fluid when it is compressed and expanded. Equation (4.5.7) describes $U$ in terms of the pressure and the equation of state.

There is an external agent $\mathcal{F}$ that changes $\mathcal{H}$. For momentum, it is the impulse of the pressure acting on the control volume, which is positive for $p\left(x_{1}\right)$ and negative for $p\left(x_{2}\right)$. The impulse is the net force times $d t$, so we define $\mathcal{F}=\left[p\left(x_{1}\right)-p\left(x_{2}\right)\right] d t$ corresponding to a unit cross section. The external agent changing the energy is the net work done by the pressure at the two ends. This work may be evaluated as the instantaneous power multiplied by $d t$, so we set $\mathcal{F}=\left[p\left(x_{1}\right) v\left(x_{1}\right)-p\left(x_{2}\right) v\left(x_{2}\right)\right] d t$. Mass is conserved, so $\mathcal{F}=0$ for conservation of mass.

The description of the contributions from the chunks of fluid that enter and leave the control volume is based on their infinitesimal extent. According to the central limit theorem, $\mathcal{H}(x, t)$ within each chunk is the same as the value at the adjacent end of the control volume. Hence, $\mathcal{H}\left(x_{1}, t\right) d s_{\text {in }}$ enters the control volume and $\mathcal{H}\left(x_{1}, t\right) d s_{\text {out }}$ exits. We also must account for the amount of the $\mathcal{H}$ quantity contained within the
control volume at instant $t$, which is obtained by integrating $\mathcal{H}(x, t)$ over $x_{1}<x<$ $x_{2}$. Let $\mathcal{I}(t)$ denote this amount.

The existence of a shock means that $\mathcal{H}$ is not a continuous function of $x$, so the integration to describe $\mathcal{I}$ must be done in a piecewise manner from $x_{1}$ to $x_{\text {dis }}$ and then from $x_{\text {dis }}$ to $x_{2}$. The positions to the left and right of the discontinuity are $x^{-}$and $x^{+}$, so the integration over the control volume at time $t$ is described by

$$
\begin{equation*}
\mathcal{I}(t) \equiv \int_{x_{1}}^{x_{2}} \mathcal{H}(x, t) d x=\int_{x_{1}}^{x^{-}} \mathcal{H}(x, t) d x+\int_{x^{+}}^{x_{2}} \mathcal{H}(x, t) d x \tag{13.2.33}
\end{equation*}
$$

We could follow similar arguments to describe the integral at time $t+d t$, with the shock situated at $x_{\text {dis }}(t+d t)$. A simpler procedure employs Leibnitz' rule for differentiating an integral with a variable limit. The difference between $x^{-}$and $x^{+}$is infinitesimal, so $d x^{-} / d t=d x^{+} / d t=d x_{\text {dis }} / d t \equiv v_{\text {dis }}$. The result is that

$$
\begin{align*}
\mathcal{I}(t+d t)= & \mathcal{I}(t)+d t \frac{d}{d t} \mathcal{I}(t) \\
= & \mathcal{I}(t)+\left\{\int_{x_{1}}^{x^{-}} \frac{\partial}{\partial t} \mathcal{H}(x, t) d x+\int_{x^{+}}^{x_{2}} \frac{\partial}{\partial t} \mathcal{H}(x, t) d x\right.  \tag{13.2.34}\\
& \left.+v_{\text {dis }}(t)\left[\mathcal{H}\left(x^{-}, t\right)-\mathcal{H}\left(x^{+}, t\right)\right]\right\} d t
\end{align*}
$$

Each balance principle states that $I(t+d t)$ must equal $I(t)$ plus the net amount of $\mathcal{H}$ that enters the control volume plus the increment due to the $\mathcal{F}$ quantity, that is,

$$
\begin{equation*}
\mathcal{I}(t+d t)=\mathcal{I}(t)+\mathcal{H}\left(x_{1}, t\right) v\left(x_{1}, t\right) d t-\mathcal{H}\left(x_{2}, t\right) v\left(x_{2}, t\right) d t+\mathcal{F} d t \tag{13.2.35}
\end{equation*}
$$

This relation applies in any situation. Substitution of the representations of $\mathcal{I}$ at $t$ and $t+d t$ into the conservation equation reduces it to

$$
\begin{gather*}
\int_{x_{1}}^{x^{-}} \frac{\partial}{\partial t} \mathcal{H}(x, t) d x+\int_{x^{+}}^{x_{2}} \frac{\partial}{\partial t} \mathcal{H}(x, t) d x+v_{\mathrm{dis}}(t)\left[\mathcal{H}\left(x^{-}, t\right)-\mathcal{H}\left(x^{+}, t\right)\right] \\
=\mathcal{H}\left(x_{1}, t\right) v\left(x_{1}, t\right)-\mathcal{H}\left(x_{2}, t\right) v\left(x_{2}, t\right)+\mathcal{F} \tag{13.2.36}
\end{gather*}
$$

The last step is to shrink the control volume by letting $x_{1}$ approach $x_{\text {dis }}$ from below, $x_{1} \rightarrow x^{-}$, and letting $x_{2}$ approach $x_{\text {dis }}$ from above, $x_{2} \rightarrow x^{+}$. This operation reduces the integrals to zero. In addition, because $x_{1}$ and $x_{2}$ now are at the shock, we have $v\left(x_{1}, t\right) \rightarrow v_{\text {dis }}$ and $v\left(x_{2}, t\right) \rightarrow v_{\text {dis }}$. To simplify the notation, let $\mathcal{H}^{-} \equiv \mathcal{H}\left(x^{-}, t\right)$ and $\mathcal{H}^{+} \equiv \mathcal{H}\left(x^{+}, t\right)$. A minor rearrangement of terms lead to a general description of each balance principle as

$$
\begin{equation*}
\mathcal{H}^{+}\left(v^{+}-v_{\mathrm{dis}}\right)=\mathcal{H}^{-}\left(v^{-}-v_{\mathrm{dis}}\right)+\mathcal{F} \tag{13.2.37}
\end{equation*}
$$

Replacing the symbols $\mathcal{H}$ and $\mathcal{F}$ with their definition for each balance principle leads to the Rankine-Hugoniot (jump) conditions. The mass equation is

$$
\begin{equation*}
\rho^{+}\left(v^{+}-v_{\mathrm{dis}}\right)=\rho^{-}\left(v^{-}-v_{\mathrm{dis}}\right) \tag{13.2.38}
\end{equation*}
$$

The momentum equation is

$$
\begin{equation*}
\rho^{+} v^{+}\left(v^{+}-v_{\mathrm{dis}}\right)+p^{+}=\rho^{-} v^{-}\left(v^{-}-v_{\mathrm{dis}}\right)+p^{-} \tag{13.2.39}
\end{equation*}
$$

The energy condition is

$$
\begin{gather*}
{\left[\frac{1}{2} \rho^{+}\left(v^{+}\right)^{2}+U^{+}\right]\left(v^{+}-v_{\mathrm{dis}}\right)+p^{+} v^{+}=\left[\frac{1}{2} \rho^{-}\left(v^{-}\right)^{2}+U^{-}\right]}  \tag{13.2.40}\\
\left(v^{-}-v_{\mathrm{dis}}\right)+p^{-} v^{-}
\end{gather*}
$$

These relations are general and therefore apply to nonacoustic applications such as blast waves and combustion in ramjet engines, as well as shocks formed by supersonic aircraft.

Our objective is to determine where the discontinuity is located and how fast it is moving. The analysis is simplified if we limit consideration to weak nonlinearity, in which case the pressure on either side of the shock is much less than the bulk modulus. Nevertheless, the operations are intricate, see for example Pierce's treatment, ${ }^{11} \mathrm{We}$ shall only take an overview of the analysis. A key step is to form a linear combination of the Rankine-Hugoniot equations in order to obtain an expression for the enthalpy per unit mass, $h=(U+p v) / \rho$. This quantity is important because the differential increment of entropy per unit mass in a process is related to the change of $h$ according to $T d S=d h-(d p) / \rho$. The order of magnitude of $p /\left(\rho_{0} c^{2}\right)$ is taken to be $O(\varepsilon)$. Simplification of the recast Rankine-Hugoniot equations based on keeping the largest terms eventually leads to recognition that the order of magnitude of $d S$ is $O\left(\varepsilon^{3}\right)$. In other words, entropy is essentially constant across a weak shock.

The implication of constancy of entropy is that the nonlinear acoustic relations derived from the Riemann solution remain valid everywhere other than $x_{\text {dis }}$. This property leads to a relation between the velocity of the shock and state of the fluid behind and ahead of the shock. Let $v_{\mathrm{av}}=\left(v^{-}+v^{+}\right) / 2$ and $\tilde{c}_{\mathrm{av}}=\tilde{c}\left(v_{\mathrm{av}}\right)$. The weakly nonlinear properties lead to

$$
\begin{equation*}
v^{-}=\frac{p^{-}}{\rho_{0} c}, \quad v^{+}=\frac{p^{+}}{\rho_{0} c}, \quad v_{\mathrm{av}}=\frac{p^{-}+p^{+}}{2 \rho_{0} c}, \quad \tilde{c}_{\mathrm{av}}=\frac{1}{2}\left[\tilde{c}\left(v^{-}\right)+\tilde{c}\left(v^{+}\right)\right] \tag{13.2.41}
\end{equation*}
$$

The velocity of the discontinuity is found to be

$$
\begin{equation*}
v_{\mathrm{dis}}=\tilde{c}_{\mathrm{av}}+v_{\mathrm{av}} \tag{13.2.42}
\end{equation*}
$$

[^72]These relations describe a situation in which the shock is propagating in the positive $x$ direction. Because a shock is characterized by a forward-leaning of the spatial profile, $p^{-}$will be greater than $p^{+}$. Thus, the shock's location will move slower than the phase speed behind the shock and faster than the phase speed ahead. If it should happen that $p^{+}=-p^{-}$, then $\tilde{c}_{\mathrm{av}}$ will be $c$. In other words, the shock will move at the linear speed of sound.

## Equal-Area Rule

It still remains to determine how $x_{\text {dis }}$ depends on time. One approach solves a differential equation based on setting $v_{\text {dis }}=d x_{\text {dis }} / d t$ in Eq. (13.2.42). Figure 13.11 describes the parameters required to carry out this analysis. In it, the multivalued pressure profile is the Riemann solution beyond the shock formation distance. Suppose the value of $x_{\text {dis }}$ is known at instant $t$. We can use the method in Sect.13.1.5 to determine the characteristics $\tau_{1}$ and $\tau_{3}$, and the associated pressures corresponding to the values of $x_{\text {dis }}$ and $t$. (The pressure at $\tau_{2}$ is irrelevant to this discussion.) The unshocked portion of the waveform adjacent to $\tau_{3}$ is farther advanced than $x_{\text {dis }}$, so $p^{+}=p\left(\tau_{3}\right)$. Similarly, the unshocked portion of the waveform adjacent to $\tau_{1}$ is behind the shock, so $p^{-}=p\left(\tau_{1}\right)$. Because $\tau_{1}$ and $\tau_{3}$ are functions of $x_{\text {dis }}$ and $t$, so too are $p^{-}$and $p^{+}$. This means that Eqs. (13.2.41) and (13.2.42) combine to give $v_{\mathrm{av}}$ as a function of $x_{\text {dis }}$ and $t$. Therefore, Eq. (13.2.42) is a statement that $d x_{\text {dis }} / d t=f\left(x_{\text {dis }}, t\right)$-it is a differential equation. The initial condition is the value of $x_{\text {dis }}$ at the instant when the profile first develops a vertical tangency.


Fig. 13.11 Relationship of the parameters for a shock to the pressure waveform obtained from the Reimann solution

Although this procedure is feasible, numerical methods would be required to solve the differential equation, and the need to find multiple roots of the Riemann solution at each integration step would require considerable effort. Furthermore, the method provides little understanding of the underlying process. Fortunately, there is an alternative that provides a graphical interpretation and determines $x_{\text {dis }}$ directly. Figure 13.12 shows a multivalued profile obtained from the Riemann solution. The difference between it and the previous figure is that the portions of this waveform that are advanced beyond $x_{\text {dis }}$ and retarded behind $x_{\text {dis }}$ are shaded. The ultimate result will be that $x_{\text {dis }}$ is such that the total shaded area, positive for the region where $x>x_{\text {dis }}$


Fig. 13.12 Spatial pressure profile obtained by evaluation of the Riemann solution beyond the location where a shock begins to form. This profile is a multivalued function of $x$ for characteristic values in the range $\tau^{-}<\tau<\tau^{+}$. The position of the area of the profile at $x>x_{\text {dis }}$ is positive, while the area for $x<x_{\text {dis }}$ is negative
and negative for $x<x_{\text {dis }}$, is zero. We shall prove this property, because doing so leads to a method by which $x_{\text {dis }}$ can be determined at any instant.

If $t$ is specified, then $p$ and $x$ are unique functions of the characteristic variable $\tau$. The values of $\tau$ corresponding to $p^{-}$and $p^{+}$are $\tau^{-}$and $\tau^{+}$, respectively. The lower portion of the profile after the discontinuity is farther advanced than the discontinuity, so $p^{+}$is the smallest value of $p$ obtained from the Riemann solution when $x=x_{\text {dis }}$ and $t$ is the time for the profile. Similarly, the upper portion of the profile preceding the discontinuity is less advanced than the discontinuity, so $p^{-}$is the largest value of $p$ obtained from the Riemann solution at $x=x_{\text {dis }}$ when $t$ is the designated time. If follows that the range of characteristic variables associated with the shaded area is $\tau^{+}<\tau<\tau^{-}$. The small-signal approximation of the equation for a characteristic tells us that the location of a specific $p(\tau)$ at instant $t$ is

$$
\begin{equation*}
x(\tau, t)=(t-\tau)(c+\beta v)=(t-\tau)\left(c+\frac{\beta}{\rho_{0} c} p(\tau)\right) \tag{13.2.43}
\end{equation*}
$$

As suggested by Fig. 13.12, the total area $\mathcal{A}$ is best evaluated by accumulating $\left(x-x_{\mathrm{dis}}\right) d p$, so that

$$
\begin{equation*}
\mathcal{A}(t)=\int_{p^{+}}^{p^{-}}\left[x(\tau, t)-x_{\mathrm{dis}}(t)\right] d p \tag{13.2.44}
\end{equation*}
$$

If $\mathcal{A}(t)$ is constant, then it should be that $d \mathcal{A} / d t=0$. Leibnitz' rule is used to differentiate the integral,

$$
\begin{align*}
\frac{d \mathcal{A}}{d t} & =\int_{p^{+}}^{p^{-}}\left[\frac{d}{d t} x(\tau, t)-v_{\mathrm{dis}}\right] d p+\left[x\left(\tau^{-}, t\right)-x_{\mathrm{dis}}\right] \frac{\partial p^{-}}{\partial t}  \tag{13.2.45}\\
& -\left[x\left(\tau^{+}, t\right)-x_{\mathrm{dis}}\right] \frac{\partial p^{+}}{\partial t}
\end{align*}
$$

The definitions of $\tau^{-}$and $\tau^{+}$are that $x\left(\tau^{-}, t\right)=x\left(\tau^{+}, t\right)=x_{\text {dis }}(t)$, so both terms obtained from the integration limits vanish. We know from the boundary condition that $p(\tau)=\epsilon \rho_{0} c V(\tau)$, so $p$ is a single-valued function of $\tau$. Accordingly, we change the integration variable from $p$ to $\tau$, so that $d p=(\partial p / \partial \tau) d \tau$. Also, differentiation of Eq. (13.2.43) with respect to $t$, with $\tau$ held fixed, gives

$$
\begin{equation*}
\frac{d x}{d t}=c+\frac{\beta}{\rho_{0} c} p(\tau) \tag{13.2.46}
\end{equation*}
$$

The result is that

$$
\begin{align*}
\frac{d \mathcal{A}}{d t} & =\int_{\tau^{+}}^{\tau^{-}}\left(c+\frac{\beta p(\tau)}{\rho_{0} c}-v_{\mathrm{dis}}\right) \frac{\partial}{\partial \tau} p(\tau) d \tau \\
& =\left.\left[\left(c-v_{\mathrm{dis}}\right) p(\tau)+\frac{\beta}{2 \rho_{0} c} p(\tau)^{2}\right]\right|_{\tau=t^{+}} ^{\tau=t^{-}} \tag{13.2.47}
\end{align*}
$$

A rearrangement of terms lead to

$$
\begin{equation*}
\frac{d \mathcal{A}}{d t}=\left(p^{-}-p^{+}\right)\left[c-v_{\mathrm{sh}}+\frac{p^{-}-p^{+}}{2 \rho_{0} c}\right] \tag{13.2.48}
\end{equation*}
$$

The term in the bracket is zero according to Eq. (13.2.42), which was derived from the Rankine-Hugoniot relations. Thus, we have proven that $d \mathcal{A} / d t=0$. This means that the total area $\mathcal{A}$ is constant at its value at the instant when the shock first formed, which is $\mathcal{A}=0$. Because the total area bounded by a graph of $p(x . t)$ versus $x$ and the line $x=x_{\text {dis }}$ is zero, then the area to the left of the discontinuity must equal the area to the right. This is the equal-area rule.

## Evaluation of Spatial Profiles

Enforcement of the equal-area rule leads to a method for determining the location of the discontinuity at any instant. There is a common situation in which little analysis is required. Let $\tau_{0}$ be a characteristic along which the pressure is zero. Suppose that $p(\tau)$ is a continuous odd function relative to $\tau_{0}$, that is, $p\left(\tau_{0}+\delta\right)=-p\left(\tau_{0}-\delta\right)$. It follows that $p^{+}=-p^{-}$regardless of $t$, so that $v_{\text {dis }}=c$ and $x_{\text {dis }}(t)=x\left(\tau_{0}, t\right)=c\left(t-\tau_{0}\right)$.

The application of the equal-area rule when $p(\tau)$ is an arbitrary function begins by selecting a trial value for the discontinuity's location; let us denote this choice as $x_{\mathrm{tr}}$. Obviously, $x_{\mathrm{tr}}$ must be somewhere in the interval where the profile is multivalued. The area $\mathcal{A}$ created in this manner will depend on $x_{\mathrm{tr}}$, as well as the value of $t$ for the profile. The basic notion is to identify which $x_{\mathrm{tr}}$ in a range of values yields $\mathcal{A}=0$.

To do so, we need a procedure for evaluating $\mathcal{A}$. To that end, the integration variable in Eq. (13.2.44) is changed from $p$ to $\tau$, which leads to

$$
\begin{equation*}
\mathcal{A}\left(x_{\mathrm{tr}}, t\right)=\int_{\tau^{+}}^{\tau^{-}}\left[x(\tau, t)-x_{\mathrm{tr}}\right] \frac{\partial}{\partial \tau} p(\tau) d \tau \tag{13.2.49}
\end{equation*}
$$

This expression can be specialized to the situation where the velocity at $x=0$ is $\varepsilon c V(t)$. Then, the weakly nonlinear pressure is $\rho_{0} c^{2} \varepsilon V(\tau)$. In view of the position dependence in Eq. (13.2.43), the area function is

$$
\begin{equation*}
\mathcal{A}\left(x_{\mathrm{tr}}, t\right)=\rho_{0} c^{2} \varepsilon \int_{\tau^{+}}^{\tau^{-}}\left\{c(t-\tau)[1+\beta \varepsilon V(\tau)]-x_{\mathrm{tr}}\right\} \dot{V}(\tau) d \tau \tag{13.2.50}
\end{equation*}
$$

Several terms are exact differentials. Integration by parts is used to handle the terms that contain $\tau$ as a factor. The result is

$$
\begin{align*}
A\left(x_{\mathrm{tr}}, t\right)= & \left.\rho_{0} c^{2} \varepsilon\left\{c\left(t-\tau-x_{\mathrm{tr}}\right) V(\tau)+\frac{1}{2} \beta \varepsilon c(t-\tau) V(\tau)^{2}\right\}\right|_{\tau^{+}} ^{\tau^{-}}  \tag{13.2.51}\\
& +\rho_{0} c^{2} \varepsilon \int_{\tau^{+}}^{\tau^{-}} c\left[V(\tau)+\frac{1}{2} \beta \varepsilon V(\tau)^{2}\right] d \tau
\end{align*}
$$

A determination of $x_{\text {dis }}$ for the spatial profile at a specified $t$ may be performed by following a set procedure. It begins by letting $x_{\mathrm{tr}}$ be the smallest $x$ in the multivalued range, solving for the smallest and largest roots, $\tau^{+}$and $\tau^{-}$, respectively, of Eq. (13.2.43) when $x=x_{\mathrm{tr}}$ and $t$ is the instant of interest, then evaluating $\mathcal{A}$ for that choice of $x_{\mathrm{tr}}$. Progressively incrementing $x_{\mathrm{tr}}$ will reduce the value of $\mathcal{A}$. The value of $x_{\text {dis }}$ lies between the values of $x_{\mathrm{tr}}$ for which $\mathcal{A}$ changes sign.

## Evaluation of Waveforms

We have established that the equal-area rule can be used to determine where a discontinuity occurs in a pressure profile at an arbitrary instant. It also can be used to determine the instant at which the discontinuity occurs in the waveform at a designated $x$. The multivalued waveform in Fig. 13.13 is the result obtained directly from the Riemann solution. The instant at which the discontinuity is observed at this location is $t_{\text {dis }}(x)$, which leads to a total area $\mathcal{A}^{\prime}\left(x, t_{\text {dis }}\right)$. The interval in which the waveform is multivalued corresponds to $p^{+}<p<p^{-}$. According to the discussion of profiles, the value of $\tau$ is a unique function of $p$, so it is best to us $d \mathcal{A}^{\prime}=\left(t_{\mathrm{dis}}-t\right) d p$. Therefore,

$$
\begin{equation*}
\mathcal{A}^{\prime}\left(x, t_{\mathrm{dis}}\right)=\int_{p^{+}}^{p^{-}}\left[t_{\mathrm{dis}}(x)-t(\tau, x)\right] d p \tag{13.2.52}
\end{equation*}
$$

To verify that the equal-area rule applies to waveforms, we evaluate $d \mathcal{A}^{\prime} / d x$. Application of Leibnitz' rule to the preceding gives

Fig. 13.13 Construction of the area bounded by a multivalued waveform and a the location of the discontinuity at an arbitrary location $x$


$$
\begin{align*}
\frac{d}{d x} \mathcal{A}^{\prime}\left(x, t_{\mathrm{dis}}\right) & =\int_{p^{-}}^{p^{+}}\left[\frac{d}{d x} t_{\mathrm{dis}}(x)-\frac{\partial}{\partial x} t(\tau, x)\right] d p+\left[t_{\mathrm{dis}}(x)-t\left(\tau^{-}, x\right)\right] \frac{\partial}{\partial x} p^{-} \\
& -\left[t_{\mathrm{dis}}(x)-t\left(\tau^{+}, x\right)\right] \frac{\partial}{\partial x} p^{+} \tag{13.2.53}
\end{align*}
$$

By definition, $t\left(\tau^{-}, x\right)=t\left(\tau^{+}, x\right)=t_{\text {dis }}(x)$, so the terms from the integration limits vanish. Also, $(d / d x) t_{\text {dis }}(x) \equiv 1 / v_{\text {dis }}(x)$. To continue the analysis, it is convenient to rewrite Eq.(13.2.43), which is the small-signal approximation of the equation for a characteristic, as

$$
\begin{equation*}
t(\tau, x)=\tau+\frac{x}{c+\beta v(\tau)} \tag{13.2.54}
\end{equation*}
$$

Differentiation of this relation with respect to $x$, with $\tau$ held fixed, yields

$$
\begin{equation*}
\frac{\partial}{\partial x} t(\tau, x)=\frac{1}{c+\beta v(\tau)} \tag{13.2.55}
\end{equation*}
$$

The result is that the spatial rate of change of $A^{\prime}$ becomes

$$
\begin{equation*}
\frac{d}{d x} \mathcal{A}^{\prime}(x)=\int_{p^{+}}^{p^{-}}\left[\frac{1}{v_{\text {dis }}(x)}-\frac{1}{c+\beta v(\tau)}\right] d p \tag{13.2.56}
\end{equation*}
$$

The analysis is restricted to weakly nonlinear signals, so $v=p /\left(\rho_{0} c\right)$ and $c \gg v(\tau)$. These approximations lead to

$$
\begin{align*}
\frac{d}{d x} \mathcal{A}^{\prime}(x) & =\int_{p^{+}}^{p^{-}}\left[\frac{c+\beta p /\left(\rho_{0} c\right)-v_{\mathrm{dis}}(x)}{c v_{\text {dis }}(x)}+\cdots\right] d p \\
& \left.=\frac{1}{c v_{\text {dis }}(x)}\left[c+\frac{1}{2} \beta \frac{p}{\rho_{0} c}-v_{\text {dis }}(x)\right]\right]\left.\right|_{p^{+}} ^{p^{-}} \tag{13.2.57}
\end{align*}
$$

The bracketed term is identically zero according to Eq. (13.2.42), so we have proven that the gradient of $A^{\prime}(x)$ is identically zero. Consequently, $\mathcal{A}^{\prime}(x)$ is constant at the value when the discontinuity began to form, so $A^{\prime}(x)=0$. Thus, we have proven that the equal-area rule applies to waveforms, as well as spatial profiles.

We may use the equal-area rule to determine when the discontinuity occurs in the waveform at a designated $x$. If the excitation function $V(\tau)$ is odd relative to a phase $\tau_{0}$ at which $V(\tau)=0$, then symmetry places the shock at $t_{\text {dis }}=\tau_{0}+x / c$. The procedure when $V(\tau)$ is arbitrary is a modification of the procedure for profiles.

To carry out the analysis, we need an expression for $\mathcal{A}^{\prime}$ in terms of the trial instant $t_{\text {tr }}$ for the discontinuity. To that end, we use the equation for a characteristic, Eq. (13.2.54), to replace $t$ in Eq. (13.2.52) and then set $v=\varepsilon c V(\tau)$ and $p=\varepsilon \rho_{0} c^{2} V(\tau)$, and change the integration variable to $\tau$. These operations give

$$
\begin{equation*}
\mathcal{A}^{\prime}\left(x, t_{\mathrm{tr}}\right)=\varepsilon \rho_{0} c^{2} \int_{\tau^{+}}^{\tau^{-}}\left[t_{\mathrm{tr}}-\tau-\frac{x}{c+\beta \varepsilon c V(\tau)}\right] \frac{d V(\tau)}{d \tau} d \tau \tag{13.2.58}
\end{equation*}
$$

A binomial series simplifies the fraction, so that

$$
\begin{equation*}
\mathcal{A}^{\prime}\left(x, t_{\mathrm{tr}}\right) \approx \varepsilon \rho_{0} c^{2} \int_{\tau^{+}}^{\tau^{-}}\left[t_{\mathrm{tr}}-\tau-\frac{x}{c}(1-\beta \varepsilon V)\right] \frac{d V(\tau)}{d \tau} d \tau \tag{13.2.59}
\end{equation*}
$$

We apply integration by parts to the term for which $\tau$ is a coefficient. All other terms are perfect differentials, so that

$$
\begin{equation*}
\mathcal{A}^{\prime}\left(x, t_{\mathrm{tr}}\right)=\rho_{0} c^{2} \varepsilon\left\{\left.\left[\left(t_{\mathrm{tr}}-\tau-\frac{x}{c}\right) V(\tau)+\frac{1}{2} \beta \varepsilon \frac{x}{c} V(\tau)^{2}\right]\right|_{\tau^{+}} ^{\tau^{-}}+\int_{\tau^{+}}^{\tau^{-}} V(\tau) d \tau\right\} \tag{13.2.60}
\end{equation*}
$$

The value of $t_{\mathrm{dis}}$ is the value of $t_{\mathrm{tr}}$ that gives $A^{\prime}=0$ when $x$ is set. Carrying out the search for this value of $t_{\mathrm{tr}}$ requires determination of the values of $\tau^{+}$and $\tau^{-}$corresponding to specified values of $x$ and $t_{\mathrm{tr}}$. These values respectively are the smallest and largest roots extracted from Eq.(13.2.54). The procedure for carrying out the analysis is described in the next example.

EXAMPLE 13.4 In Example 13.2, the excitation at $x=0$ is a harmonically varying velocity starting at $t=0$. Determine the waveforms at $x /\left(x_{\text {shock }}\right)_{\text {sine }}=$ $1.2,1.5,2$, and 4.

## Significance

A primary focus is development of a general computational algorithm for fitting shocks into the Riemann solution. The resulting waveforms will shed light on the trend for the signal to approach a sawtooth waveform.

## Solution

The starting point for the analysis is the evaluation of the Riemann solution. The direct computational method used to solve Example 13.2 may be used for this purpose because there is no need to have the pressure data on a base of equally spaced $t$ values. The nondimensional shock formation distance is $k x_{\text {shock }}=1 /(\varepsilon \beta)$, so we seek the waveform at locations where $k x=\sigma k x_{\text {sh }}=\sigma /(\beta \varepsilon)$. The acoustic Mach number is $\varepsilon=0.002$. The Riemann solution at a set of equally spaced $\omega \tau_{n}$ values starting at $\omega \tau_{1}=0$ yields a set of $q_{n} \equiv p /\left(\rho_{0} c^{2}\right)$ values corresponding to instants $\omega t_{n}$, where

$$
\begin{gather*}
q_{n}=\varepsilon \sin \left(\omega \tau_{n}\right) h\left(\omega \tau_{n}\right)  \tag{1}\\
\omega t_{n}=\omega \tau_{n}+\frac{\sigma}{\beta \varepsilon} \frac{1}{1+\beta \varepsilon \sin \left(\omega \tau_{n}\right) h\left(\omega \tau_{n}\right)}
\end{gather*}
$$

The evaluation of the Riemann solution yields a correlated set of $\omega \tau_{n}, \omega t_{n}$, and $q_{n}$ data at fixed $\sigma$ in which the constant interval for the $\omega \tau$ values is $2 \pi / N$.

The Riemann solution beyond $x_{\text {shock }}$ is typified by the waveform at $\sigma=4$ in Fig. 1. Two different conditions are seen to exist. At the leading edge, which corresponds to $\tau=0$, there is no prior signal. The multivalued waveform has no special symmetry properties in this interval, so fitting a discontinuity will require an analysis of the waveform area. Subsequent to arrival of the shock at the leading edge, the multivalued intervals are centered on $\omega \tau=2 j \pi$, which is the phase at which $p=0$ and $d p / d \tau$ is the maximum positive value. With respect to these instants, $p$ is an odd function. Thus, the shock will occur at the time values for which $\omega\left(t_{\text {dis }}-x / c\right)=2 j \pi$.


Figure 1.

We begin by modifying the computed waveform in the periods after the first. Let $\tau_{J}$ be any instant for which $p\left(\tau_{J}\right)=0$. In view of the periodic nature of $p(\tau)$, and the fact that $p(0)=0$, the second and subsequent discontinuities correspond to $J=$ $N+1,2 N+1, \ldots$. Therefore, $\omega t_{\text {dis }}=\omega \tau_{J}+k x$ for these discontinuities. Now that
we know the instant at which a discontinuity occurs, we can correct the multivalued pressure. Smaller $\tau$ values that occur later than $\omega t_{J}$, and subsequent $\tau$ values that occur earlier are physically impossible. These phases belong to the discontinuity. This condition can occur for no more than $N / 2$ points on either side of $\omega \tau_{J}$. Thus, the procedure is to set $\omega \tau_{J-n}=\omega \tau_{J}, \omega t_{J-n}=\omega t_{\text {dis }}$, and $p_{J-n}=0$ if $\omega \tau_{J-n}>\omega \tau_{J}$, and $\omega \tau_{J+n}=\omega \tau_{J}, \omega t_{J-n}=\omega t_{\text {dis }}$, and $p_{J+n}=0$ if $\omega \tau_{J+n}<\omega \tau_{J}$. These operations are implemented for $n=1, \ldots, N / 2$.

Fitting the shock at the leading edge requires explicit examination of the waveform area according to Eq. (13.2.60) for each of the Riemann data values. The excitation function is $V(\tau)=\sin (\omega \tau) h(\omega t)$. Furthermore, the area of the waveform is situated in $p \geq 0$, so the lower limit in that expression corresponds to $\tau^{+}=0$ and $V\left(\tau^{+}\right)=0$. Thus, the waveform area corresponding to a trial instant $\omega t_{\mathrm{tr}}$ for the discontinuity is

$$
\begin{align*}
\omega \mathcal{A}^{\prime}\left(k x, \omega t_{\mathrm{tr}}\right) & =\rho_{0} c^{2} \varepsilon\left\{\left[\left(\omega t_{\mathrm{tr}}-\omega \tau^{-}-k x\right) \sin \left(\omega \tau^{-}\right)+\frac{1}{2} \beta \varepsilon k x \sin \left(\omega \tau^{-}\right)^{2}\right]\right. \\
& \left.+\left[1-\cos \left(\omega \tau^{-}\right)\right]\right\} \tag{1}
\end{align*}
$$

The true value of $\omega t_{\text {dis }}$ is obtained by setting $\omega t_{\mathrm{tr}}=\omega t_{n}$ for a range of $n$ and then searching for the value of $n$ that best fits the criterion that $\omega \mathcal{A}^{\prime}\left(k x, \omega t_{n}\right)=0$. Explanation of the procedure is facilitated by considering Fig. 2, which zooms in on the multivalued interval at the leading edge. The dots mark points on the waveform that are the computed data. The first point, $\omega \tau_{1}=0$, corresponds to the largest $t$ in the multivalued interval. The earliest time in the data set is $\omega t_{M}$. The value of $M$ is determined by seeking the index corresponding the minimum of the $\omega t_{n}$ data.


Figure 2.

The search procedure entails setting $\omega t_{\text {tr }}$ equal to a computed $\omega t_{n}$ value in the range $1 \leq n \leq M$. By definition, $\omega \tau^{-}$is the characteristic value on the upper part of the multivalued segment when the time is $\omega t_{\mathrm{tr}}$. In Fig. 2, this point is the intersection of the vertical line through point $n$ and the upper branch. Because a constant increment of the $\omega \tau$ values leads to uneven spacing of the $\omega t$ values, this intersection will not fall on a point that was computed along the upper branch. Therefore, the value of $\omega \tau^{-}$is not yet known, but evaluation of the area in Eq. (1) requires that we locate this intersection. There are several options for this determination. The most accurate one uses the method in Sect. 13.1.5 to determine the largest of the three values of $p$
given by the Riemann solution at $k x$ and this $\omega t_{\mathrm{tr}}$. A simpler alternative interpolates between data points on the upper branch on either side of $\omega t_{\mathrm{tr}}$. The simplest to implement selects the computed point on the upper branch that is closest to $\omega t_{n}$. The latter procedure will be adequate if the value of $N$ is sufficiently large. Regardless of which method is used, the concept is to use the value of $\omega \tau^{-}$to compute $\omega \mathcal{A}^{\prime}$. This is done beginning at $n=1$, and ending at the most at $n=M$. The search may be halted when the sign of $\mathcal{A}^{\prime}$ changes.

The resulting values of $\omega \mathcal{A}^{\prime}$ in the full range of possible $\omega t_{\text {tr }}$ are shown in Fig. 3 for $x=1.2 x_{\mathrm{sh}}$. The value of $\mathcal{A}^{\prime}$ is zero at $\omega t_{\mathrm{tr}}=199.9464$, so this is the value of $\omega t_{\mathrm{dis}}$. The computations used $N=512$, which corresponds to $\Delta(\omega \tau)=2 \pi / 512$. This is a reasonably fine resolution, but the nearly linear nature of the plotted curve suggests that it would have been adequate to use a smaller $N$.


Figure 3.

After $\omega t_{\text {dis }}$ has been identified, the waveform may be corrected. The first step is to identify the instant that is closest to $\omega t_{\text {dis }}$ on the middle branch. We denote this instant as $\omega t_{K}$. The index of points that lie on the middle branch is $1<K<M$. The points in the multivalued interval that are delayed too much are adjusted by assigning them to the bottom of the discontinuity, so that $\omega t_{n}=\omega t_{\text {dis }}$ and $p_{n}=0$ if $\omega t_{n}>\omega t_{\text {dis }}$ and $n<K$. Data values that occur too early are shifted to the top of the discontinuity by setting $\omega t_{n}=\omega t_{\text {dis }}$ and $p_{n}=p_{K}$ if $\omega t_{n}<\omega t_{\text {dis }}$ and $n>K$.

These procedures yield a set of $(\omega t, p)$ data that is single-valued function of time. Figure 4 is the result at the four specified locations. At $20 \%$ farther than the location where the shock begins to occur, the discontinuity is barely noticeable. Nevertheless, there is a visible similarity of the waveform to the teeth of a saw. Propagation to farther distances enhances this appearance. Indeed, at $\sigma=4$, there is no evidence of a sinusoidal function. Another important aspect is the effect of the discontinuity on the peak pressure. At $\sigma=1.2$ and $\sigma=1.5$, the maximum pressure in the Riemann solution occurs slightly after each of the multivalued intervals. In contrast, at $\sigma=2$ and at $\sigma=4$ the maximum and minimum pressures in the Riemann solution occur in the multivalued interval. Thus, fitting a discontinuity to those waveforms results in reduction of the peak pressure. From this, we may deduce that the shock wave dissipates energy.


Figure 4.

An interesting question is how the discontinuities propagate relative to the waveform. The waveform is an odd function relative to the second and subsequent discontinuities. These discontinuities propagate at the linear speed of sound, so their relative spacing is maintained at $c T$, and the interval between their occurrence in a waveform will be $T$. This is not true for the leading discontinuity because for it, $p^{-}$ is positive and $p^{+}=0$. The average pressure across this shock is $p_{\text {av }}=p^{-} / 2$. The Rankine-Hugoniot relation, Eq. (13.2.42), states that this discontinuity will propagate at $\tilde{c}\left(p^{-} /\left(2 \rho_{0} c\right)\right)+p^{-} /\left(2 \rho_{0} c\right)>c$. Therefore, the distance from the shock at the leading edge to the next shock will decrease with increasing $t$. For the same reason, the leading edge shock will precede arrival of the next shock by more than $T$. Eventually, the value of $p^{-}$progressively decreases as the discontinuity cuts off an increasing portion of the multivalued waveform. In terms of nondimensional variables, the discontinuity's arrival at a specified $x$ is advanced by $\omega t_{\text {dis }}-k x$. Figure 5 is the result of computing this quantity over a broad range of $\sigma$ values. Plotting the time advancement as a ratio to $2 \pi$ allows us to see it as a fraction of a period. The rate of this advancement decreases with increasing distance due to the decrease of $p^{-}$. These features are examined more closely in the next section.


Figure 5.

## Old-Age Behavior

The previous example demonstrates the drastic ways that a signal is affected by the creation of shocks. These effects grow as a signal propagates well beyond the location where a shock first forms. A square wave provides a convenient means for exploring these phenomena because the equal-area rule can be enforced without requiring integration to evaluate the waveform area. Although this is the only type of waveform we will consider, the observations are widely applicable because of the tendency of all waves to evolve to a sawtooth appearance.

The excitation is a particle velocity at $x=0$ that is a periodic square wave beginning at $t=0$. The peak particle velocity is set at $\varepsilon c$, with $\varepsilon \ll 1$. The small-signal approximation tells us that $p=\rho_{0} c v$, so the pressure also is a square wave, as shown in Fig. 13.14, where the maximum value is $\hat{p}=\rho_{0} c^{2} \varepsilon$. The presence of discontinuities in which the pressure increases tells us that the signal at $x=0$ already contains shocks.


Fig. 13.14 A periodic square wave generated at $x=0$. The pressure is zero for $t<0$

The positive pressures propagate faster and the negative pressures propagate slower than the linear speed of sound. Because $|\varepsilon| \ll 1$, the Riemann solution has the property that straight segments of the waveform remain straight as the wave propagates. Thus, according to the Riemann solution, the square lobes become trapezoids that lean forward in time as the signal propagates. Figure 13.15 shows this waveform at a propagation distance that is not too large. The phase speed of the maximum positive pressures is $c(1+\beta \varepsilon)$, whereas the phase speed at zero pressure is $c$. Thus,


Fig. 13.15 Comparison of the Riemann solution and linear theory for propagation of a periodic square wave. The location is not too distant from the source
the time advancement $\Delta t$ of the positive pressure plateaus relative to the arrival time of the preceding $p=0$ phases is

$$
\begin{equation*}
\Delta t=\frac{x}{c}-\frac{x}{c(1+\beta \varepsilon)} \approx \beta \varepsilon \frac{x}{c} \tag{13.2.61}
\end{equation*}
$$

The phase speed of the negative pressure plateaus is $c(1-\beta \varepsilon)$. To $O(\varepsilon)$, these signals are retarded in time relative to $p=0$ by the same $\Delta t$. As part of the analysis, we will identify the largest distance, labeled $x_{\text {old }}$, for which this picture is appropriate.

According to the Riemann solution, $p=0$ and $\partial p / \partial \tau>0$ when $t$ equals $x / c$ plus an integer multiple of the period. The multivalued portion of the Riemann solution contains these instants. They are corrected by application of the equal-area rule.

The piecewise linear nature of the Riemann waveform simplifies this adjustment. After the first cycle, the waveform in the multivalued intervals are odd functions relative to the instant when $p=0$. Figure 13.16 shows a vertical line through $p=0$ within these intervals. The waveform area for negative $p$ is the negative of the area for positive $p$. It follows that these lines define the time $t_{\text {dis }}$ of discontinuities. The discontinuity values are $p^{-}=\hat{p}$ and $p^{-}=-\hat{p}$. The Rankine-Hugoniot relation states that shocks propagate at the linear speed of sound plus the average of the particle velocities ahead of and behind the discontinuity. This construction in Fig. 13.16 is consistent with this principle.


Fig. 13.16 Application of the equal-area rule to a periodic square wave. The propagation distance is relatively small

The situation at the leading edge is different, because no signal precedes the $p=0$ phase. A vertical line that splits $\Delta t$ leads to a waveform area that consists of identical triangles to the left and right, so the discontinuity at the leading edge is advanced by $\Delta t / 2$. It follows that the leading edge arrives at position $x$ when $t=$ $x / c-\Delta t / 2=(x / c)(1-\beta \varepsilon / 2)$. The corresponding propagation speed of this shock is $x / t=c /(1-\beta \varepsilon / 2) \approx c(1+\beta \varepsilon / 2)$. This too is consistent with the RankineHugoniot relations because it is the mean of the sound speeds for $p^{-}=\hat{p}$ and $p^{+}=0$. Increasing the propagation distance leads to an increase of $\Delta t$. At some location, the Riemann solution predicts that the last pressure on a positive plateau occurs at the instant of a discontinuity, and that the first pressure on a negative plateau also occurs at that instant. This condition, which is characterized by $\Delta t=T / 2$, is described by Fig. 13.17. Only the leading edge is exempt from this behavior.

Beyond the location described by Fig. 13.17, the waveform is altered in a fundamentally different manner that is ongoing. This position marks the advent of the old-age stage, so it is labeled as $x_{\text {old }}$. Setting $\Delta t=T / 2$ gives


Fig. 13.17 Distortion of an initial square wave at the onset of old age. The time advancement of the positive pressure equals half a period

$$
\begin{equation*}
x_{\text {old }}=\frac{c T}{2 \beta \varepsilon}=\frac{c T}{2 \beta}\left(\frac{\rho_{0} c^{2}}{\hat{p}}\right) \tag{13.2.62}
\end{equation*}
$$

In the present case of a periodic square wave, the old-age stage begins when a plateau has advanced sufficiently that its entire occurrence falls in the interval where the Riemann solution is multivalued. Beyond that occurrence, the discontinuity intersects the Riemann solution on a sloping portion. A negative plateau also is cut off because it is delayed into the multivalued interval. This is the situation in Fig. 13.18. The equal-area rule maintains all discontinuities, other than the leading edge, at the location where the Riemann solution gives $p=0$. The consequence is that the peak maximum and minimum pressure amplitudes are reduced. Any wave that propagates beyond its shock formation distance will ultimately exhibit this phenomenon. It is the hallmark of the old-age stage.


Fig. 13.18 Waveform of an initially square wave at old-age. Distortion results in a decrease of the peak pressure

An expression for the peak pressure in the old-age stage is obtained from similar triangles formed by a discontinuity and the Riemann solution, specifically

$$
\begin{equation*}
\frac{p_{\max }}{T / 2}=\frac{\hat{p}}{\Delta t} \tag{13.2.63}
\end{equation*}
$$

Substitution of $\Delta t$ from Eq. (13.2.61) and $\hat{p}=\rho_{0} c^{2} \varepsilon$ leads to

$$
\begin{equation*}
p_{\max }=\frac{\rho_{0} c^{2}}{2 \beta}\left(\frac{c T}{x}\right)=\hat{p} \frac{x_{\text {old }}}{x}, x>x_{\text {old }} \tag{13.2.64}
\end{equation*}
$$

The first form is remarkable because it indicates that the peak pressure is independent of $\hat{p}$. In other words, if all other parameters are held fixed, increasing the amplitude of the input will not alter the pressure observed at a specified location in the old-age region. Why is this so? If we increase the amplitude, the value of $x_{\text {old }}$ decreases. Consequently, the signal enters old-age at a closer location, so $x_{\text {old }} / x$ decreases. Thus, the increase in the input amplitude is offset by the greater old-age reduction. This phenomenon is called saturation. It limits the loudness of the sound that can be heard. To see why, suppose we select a location $x$. At very small values of $\hat{p}$, this location is much smaller than the value of $x_{\text {old }}$ given by Eq. (13.2.62). Increasing the amplitude of the input proportionally raises the amplitude of the Riemann solution. However, the increase also reduces the value of $x_{\text {old }}$. Further increase eventually reduces $x_{\text {old }}$ to $x$ at the field point, which occurs when $\hat{p}=\rho_{0} c^{2}(c T) /(2 \beta x)$. This is the saturation pressure. Further increase of $\hat{p}$ leads to $x>x_{\text {old }}$, which means that Eq. (13.2.64) now describes the peak pressure. As an example, consider a signal in air for which $1 / T=1 \mathrm{kHz}$. The saturation amplitude at $x=100 \mathrm{~m}$ is 197 Pa . The RMS pressure in one cycle of the sawtooth waveform is $p_{\max } / 3^{1 / 2}$, so the sound pressure level at saturation is 135 dB . The only way in which a louder sound at that location can be obtained with a plane wave is to raise $T$, that is, lower the fundamental frequency.

The reduction of peak pressure in the old-age stage brings to the fore the issue of mechanical energy. Because we are dealing with simple plane waves, and the particle velocity is $p /\left(\rho_{0} c\right)$ in the small-signal approximation, the time-averaged intensity has the same form as it does for a linear signal whose waveform is arbitrary. For the present purpose, we may use any period subsequent to the first, so that

$$
\begin{equation*}
I_{\mathrm{av}}(x)=\frac{1}{T} \int_{x / c+n T}^{x / c+(n+1) T} \frac{p(x, t)^{2}}{\rho_{0} c} d t \tag{13.2.65}
\end{equation*}
$$

For the initial waveform in Fig. 13.14, we have $I_{\text {av }}(0)=\hat{p}^{2} /\left(\rho_{0} c\right)$. The area under the square of trapezoidal waveform in Fig. 13.15 may be found by adding the contribution of the plateau and the linear decrease, which gives

$$
\begin{equation*}
I_{\mathrm{av}}(x)=\frac{2}{\rho_{0} c T}\left[\hat{p}^{2}\left(\frac{T}{2}-\Delta t\right)+\frac{1}{3} \hat{p}^{2} \Delta t\right]=\frac{\hat{p}^{2}}{\rho_{0} c}\left(1-\frac{2}{3} \frac{x}{x_{\mathrm{old}}}\right), x<x_{\mathrm{old}} \tag{13.2.66}
\end{equation*}
$$

In the region where old-age behavior has taken over, integration of the square of the triangular waveform in Fig. 13.18 leads to

$$
\begin{equation*}
I_{\mathrm{av}}(x)=\frac{2}{\rho_{0} c T}\left(\frac{1}{3} p_{\max }^{2} \frac{T}{2}\right)=\frac{1}{3} \frac{\hat{p}^{2}}{\rho_{0} c}\left(\frac{x_{\mathrm{old}}}{x}\right)^{2}, x>x_{\mathrm{old}} \tag{13.2.67}
\end{equation*}
$$

At $x_{\text {old }}$, both expressions give $I_{\text {av }}=(\hat{p})^{2} /\left(3 \rho_{0} c\right)$. We see that the time-averaged intensity decreases with increasing $x$ at all locations. This means that some energy is lost, rather than flowing downstream. Before the onset of old-age, the loss is linear. In old-age, the power flow decreases as the inverse square of distance, so the energy loss is much more drastic.

The loss of energy is a consequence of the thermodynamic mechanisms captured by the Rankine-Hugoniot relations. This effect arises even if nonideal effects are negligible. However, dissipative effects like viscosity and molecular relaxation are always present. Attenuation of harmonics composing a periodic signal typically increases with increasing harmonic number, as will be seen in Sect.13.3.2. This effect is not necessarily large, depending on the properties of the fluid. However, it combines with the old-age behavior, which reduces all harmonics. Ultimately, what remains is a fundamental harmonic that has been reduced sufficiently to permit the usage of linear acoustic theory.

One of the surprising attributes of shocks is that they affect a pulse differently from a steady-state signal. We encountered this difference in the case of continuous square wave excitation, where the discontinuity at the leading edge was unlike those embedded in the subsequent response. To explore this further, we consider a pulse at $x=0$ that is the first cycle of the square wave, after which it is cut off. In Fig. 13.19, the pulse has been propagating for a while, but it has not yet reached old-age.


Fig. 13.19 Early stage in the propagation of single period of a square wave

The Riemann solution for the pulse is the same as its prediction for the first period of the continuous wave. As a consequence, the shock at the leading edge behaves in the same way as the shock for the continuous wave. Termination of the waveform after one cycle leads to $p^{+}=-\hat{p}$ and $p^{-}=0$. The multivalued waveform at the trailing edge is not what it was for a continuous wave. Rather, it is an image of the waveform at the leading edge. Therefore, the retardation of the discontinuity at the trailing edge is the same as the advancement at the leading edge. The consequence is that the pulse duration is extended to $x / c-\Delta t / 2<t<x / c+T+\Delta t / 2$, whereas it is $x / c<t<x / c+T$ in linear theory.

Further propagation eventually brings the pulse to the transition where old-age behavior sets in. This occurs when the last pressure on the positive plateau catches up to the discontinuity, which is the same instant as that at which the first pressure on the negative plateau is delayed to the trailing discontinuity. This occurs when $\Delta t=T$, which leads to $x_{\text {old }}$ being twice as distant as the value in Eq. (13.2.62). Figure 13.20 shows the waveform beyond the old-age location.


Fig. 13.20 Distortion of a square wave pulse beyond the distance where old-age behavior begins

Application of the equal-area to locate the discontinuity in the old-age signal would be a little more difficult than it was for the continuous wave because the waveform area has a different shape on either side of the discontinuity. Nevertheless, it is evident that the waveform preserves its sawtooth shape with an ever decreasing amplitude, and that its duration increases progressively as the propagation distance increases.

It is reasonable to ponder whether the old-age behavior of a square wave pulse is representative of general trends. Evidence that it may be found in the results of Example 13.4, where the waveform strongly resembles a sawtooth at a distance as small as $x=2 x_{\text {shock. }}$. Indeed, if we set $T=2 \pi / \omega$ for the fundamental frequency, the expression in Eq. (13.2.62) gives $x_{\text {old }}=\pi\left(x_{\text {shock }}\right)_{\text {sine }}$. A comparison of the waveforms at $x=2 x_{\text {shock }}$ and $x=4 x_{\text {shock }}$ in Fig. 3 of the aforementioned example shows that oldage amplitude reduction has occurred at the larger distance. The old-age behavior of any signal will be like that of the square wave, but the details of where it sets in, where the discontinuities are situated, and how the amplitude decreases with increasing distance will depend on the nature of the input function.

### 13.3 General Analytical Techniques

We began our study of linear acoustics with the d'Alembert solution. Like the Riemann solution, it is obtained by analyzing the characteristics. To extend the scope of linear systems that could be investigated, we invoked other methods to solve the wave equation. In the same way, if we are to expand our exploration of nonlinear acoustic waves, we will need an appropriate field equation. The nonlinear wave equation we will derive contains no approximations beyond the assumption of an isentropic process in an inviscid fluid. To the author's knowledge, the first appearance of this equation was in a text on aeroelasticity. ${ }^{12}$ The derivation parallels that in a classic fluid mechanics treatise, ${ }^{13}$ and an alternative derivation appears in an acoustics text. ${ }^{14}$ After the field equation is derived, two general methods for solving it will be derived for the case of plane waves.

[^73]
### 13.3.1 A Nonlinear Wave Equation

The basic state variables are the acoustic pressure $p$, the particle velocity $\bar{v}$, and the current density $\rho$. The ambient state corresponds to $p=0, \bar{v}=\overline{0}$, and $\rho=\rho_{0}$. The field equation we seek governs the velocity potential $\phi(\bar{x}, t)$, whose gradient is the particle velocity,

$$
\begin{equation*}
\bar{v}=\nabla \phi \tag{13.3.1}
\end{equation*}
$$

In terms of this function, the continuity equation, Eq. (4.1.5), is

$$
\begin{equation*}
\frac{1}{\rho} \frac{\partial \rho}{\partial t}+\frac{\nabla \rho}{\rho} \cdot \nabla \phi+\nabla^{2} \phi=0 \tag{13.3.2}
\end{equation*}
$$

and the momentum equation, Eq. (4.1.7), is

$$
\begin{equation*}
\frac{\partial}{\partial t} \nabla \phi+\frac{1}{2} \nabla(\nabla \phi \cdot \nabla \phi)=-\frac{\nabla p}{\rho} \tag{13.3.3}
\end{equation*}
$$

The overall task now is to replace terms that contain $p$ or $\rho$ with terms that solely contain $\phi$. Our strategy entails manipulating the momentum equation and the equation of state to obtain suitable representations of the first two terms in the continuity equation. Removal of the second term is the more direct. The corollary of the assumption of constant entropy is that $p$ is a known function of $\rho$. Taking the gradient of this function shows that $\nabla p$ is proportional to $\nabla \rho$, specifically,

$$
\begin{equation*}
\nabla p=\tilde{c}^{2} \nabla \rho, \quad \tilde{c}=\left(\frac{d p}{d \rho}\right)^{1 / 2} \tag{13.3.4}
\end{equation*}
$$

The quantity $\tilde{c}$ is the tangential sound speed we first encountered in the derivation of the Riemann solution. The result of substituting this definition into the momentum equation is an expression we will use to replace the second term in the continuity equation, Eq. (13.3.2), specifically,

$$
\begin{equation*}
\frac{\nabla \rho}{\rho}=-\frac{1}{\tilde{c}^{2}}\left[\frac{\partial}{\partial t} \nabla \phi+\frac{1}{2} \nabla(\nabla \phi \cdot \nabla \phi)\right] \tag{13.3.5}
\end{equation*}
$$

The task now is to find a suitable description, the first term in the continuity equation using relations derived from the momentum equation and/or equation of state. The means for doing so entails working with the pressure, rather than the density, using an ingenious application of Leibnitz rule for differentiating an integral. According to it

$$
\begin{align*}
\frac{1}{\rho} \frac{\partial p}{\partial t} & \equiv \frac{\partial}{\partial t} \int_{0}^{p} \frac{d p^{\prime}}{\rho\left(p^{\prime}\right)} \\
\frac{\nabla p}{\rho} & \equiv \nabla \int_{0}^{p} \frac{d p^{\prime}}{\rho\left(p^{\prime}\right)} \tag{13.3.6}
\end{align*}
$$

Both identities follow from the fact that only the upper limit of the integral depends explicitly on position and time.

These relations merely replace one form of dependence on $\rho$ with another. The reason they are useful is that the right sides are respectively the time derivative or gradient of a quantity. Substitution of the gradient form into the original momentum equation, Eq.(13.3.3), gives

$$
\begin{equation*}
\nabla\left(\frac{\partial \phi}{\partial t}+\frac{1}{2} \nabla \phi \cdot \nabla \phi+\int_{0}^{p} \frac{d p^{\prime}}{\rho\left(p^{\prime}\right)}\right)=0 \tag{13.3.7}
\end{equation*}
$$

This equation requires that the term within the parentheses be independent of position, but it could be a function of time. However, it was stipulated that the fluid is initially at rest in the ambient state. The term in parentheses is zero at that initial state and therefore always. Thus, we have

$$
\begin{equation*}
\frac{\partial \phi}{\partial t}+\frac{1}{2} \nabla \phi \cdot \nabla \phi+\int_{0}^{p} \frac{d p^{\prime}}{\rho\left(p^{\prime}\right)}=0 \tag{13.3.8}
\end{equation*}
$$

In later developments, this equation will lead to a relation for the pressure corresponding to a known solution for the velocity potential. Here, we use it to eliminate the density from the continuity equation. Toward that objective, we differentiate the equation with respect to time and then recall the first of Eq. (13.3.6). The result is

$$
\begin{equation*}
\frac{\partial^{2} \phi}{\partial t^{2}}+\frac{\partial}{\partial t}\left(\frac{1}{2} \nabla \phi \cdot \nabla \phi\right)+\frac{1}{\rho} \frac{\partial p}{\partial t}=0 \tag{13.3.9}
\end{equation*}
$$

The pressure is eliminated by applying the chain rule for differentiation to the equation of state, which gives $\partial p / \partial t=\tilde{c}^{2} \partial \rho / \partial t$. Thus, these manipulations of the momentum equation have led us to a relation that matches the first density term in Eq. (13.3.2),

$$
\begin{equation*}
\frac{1}{\rho} \frac{\partial \rho}{\partial t}=-\frac{1}{\tilde{c}^{2}}\left[\frac{\partial^{2} \phi}{\partial t^{2}}+\frac{\partial}{\partial t}\left(\frac{1}{2} \nabla \phi \cdot \nabla \phi\right)\right] \tag{13.3.10}
\end{equation*}
$$

The result of substituting this relation and Eq. (13.3.5) into Eq. (13.3.2), followed by collection and rearrangement of some terms, is

$$
\begin{equation*}
\tilde{c}^{2} \nabla^{2} \phi-\frac{\partial^{2} \phi}{\partial t^{2}}-\frac{\partial}{\partial t}(\nabla \phi \cdot \nabla \phi)-\frac{1}{2} \nabla(\nabla \phi \cdot \nabla \phi)=0 \tag{13.3.11}
\end{equation*}
$$

All product terms in the preceding are the result of kinematical nonlinearity. Constitutive nonlinearity is contained in the equation of state, but that effect is buried in $\tilde{c}^{2}$. Differentiation of the equation of state will give $\tilde{c}^{2}$ as a function of either $p$ or $\rho$, depending on how the equation of state is written. Unfortunately, neither form is acceptable because we seek a field equation in which $\phi$ is the sole dependent variable.

Thus, we cannot proceed until we specify the equation of state. Let us begin by considering an ideal gas. For an isentropic process, the relation is

$$
\begin{equation*}
\frac{\rho}{\rho_{0}}=\left(\frac{p}{p_{0}}+1\right)^{1 / \gamma} \tag{13.3.12}
\end{equation*}
$$

The definition of $\tilde{c}$ gives

$$
\begin{equation*}
\tilde{c}^{2} \equiv \frac{1}{d \rho / d p}=c^{2}\left(\frac{p}{p_{0}}+1\right)^{(\gamma-1) / \gamma} \tag{13.3.13}
\end{equation*}
$$

We wish to eliminate $p$ in favor of $\phi$ in this equation. The means for doing so is to substitute the equation of state into Eq. (13.3.8). This operation makes it possible to evaluate the integral, specifically

$$
\begin{align*}
& \frac{\partial \phi}{\partial t}+\frac{1}{2} \nabla \phi \cdot \nabla \phi+\frac{1}{\rho_{0}} \int_{0}^{p}\left(\frac{p^{\prime}}{p_{0}}+1\right)^{-1 / \gamma} d p^{\prime} \\
& \equiv \frac{\partial \phi}{\partial t}+\frac{1}{2} \nabla \phi \cdot \nabla \phi+\left(\frac{c^{2}}{\gamma-1}\right)\left[\left(\frac{p}{p_{0}}+1\right)^{(\gamma-1) / \gamma}-1\right]=0 \tag{13.3.14}
\end{align*}
$$

where the relation for the linear speed of sound, $c^{2}=\gamma p_{0} / \rho_{0}$, has been used. Solving this equation for $p$ in terms of $\phi$ yields

$$
\begin{equation*}
\frac{p}{p_{0}}=\left[1-\frac{\gamma-1}{c^{2}}\left(\frac{\partial \phi}{\partial t}+\frac{1}{2} \nabla \phi \cdot \nabla \phi\right)\right]^{\gamma /(\gamma-1)}-1 \tag{13.3.15}
\end{equation*}
$$

We seek an expression for $\tilde{c}^{2}$ in terms of $p$, so we substitute the above expression for $p$ into Eq. (13.3.13). The result is

$$
\begin{equation*}
\tilde{c}^{2}=c^{2}-(\gamma-1)\left(\frac{\partial \phi}{\partial t}+\frac{1}{2} \nabla \phi \cdot \nabla \phi\right) \tag{13.3.16}
\end{equation*}
$$

Substitution of $\tilde{c}^{2}$ into Eq. (13.3.11) yields a nonlinear field equation,

$$
\begin{align*}
c^{2} \nabla^{2} \phi & -\frac{\partial^{2} \phi}{\partial t^{2}}-\left[(\gamma-1) \nabla^{2} \phi \frac{\partial \phi}{\partial t}+\frac{\partial}{\partial t}(\nabla \phi \cdot \nabla \phi)\right.  \tag{13.3.17}\\
& \left.+\frac{(\gamma-1)}{2}(\nabla \phi \cdot \nabla \phi) \nabla^{2} \phi+\frac{1}{2} \nabla \phi \cdot \nabla(\nabla \phi \cdot \nabla \phi)\right]=0
\end{align*}
$$

This is the nonlinear wave equation. When the terms contained within the bracket are dropped, what remains is the linear wave equation. The full equation governs any wave in an ideal gas. Its appearance is quite daunting. We may simplify it somewhat by invoking the small-signal approximations. The implication of $|p|$ being much less
than the bulk modulus is that $\phi$ also is small. If that is so, then the cubic terms are much less than the quadratic terms, which are much less than the linear terms. It is necessary to retain the quadratic products as the first approximation of nonlinear effects, but the cubic terms may be dropped.

In addition to simplification of the field equation, dropping the cubic terms has another beneficial feature. When the equation of state for an ideal gas is truncated at quadratic terms, it is the same as the equation for a liquid if we state both in terms of the coefficient of nonlinearity in Eq. (13.1.34). Hence, setting $\gamma+1=2 \beta$ in the nonlinear wave equation and deleting the cubic terms yield a field equation that is descriptive of weakly nonlinear waves in an ideal gas, as well as almost any wave in a liquid. This version of the nonlinear wave equation is

$$
\begin{equation*}
c^{2} \nabla^{2} \phi-\frac{\partial^{2} \phi}{\partial t^{2}}-\left[2(\beta-1) \nabla^{2} \phi \frac{\partial \phi}{\partial t}+\frac{\partial}{\partial t}(\nabla \phi \cdot \nabla \phi)\right]=0 \tag{13.3.18}
\end{equation*}
$$

We will seek solutions of this equation. However, the state variable of interest is the pressure. Equation (13.3.15) describes $p$ in terms of $\phi$. This relation also is simplified by the small-signal approximations. The bracketed term is expanded in a binomial series. Truncation of the expansion at quadratic terms gives

$$
\begin{equation*}
\frac{p}{p_{0}}=-\frac{\gamma}{c^{2}}\left[\frac{\partial \phi}{\partial t}+\frac{1}{2} \nabla \phi \cdot \nabla \phi-\frac{1}{2 c^{2}}\left(\frac{\partial \phi}{\partial t}\right)^{2}\right] \tag{13.3.19}
\end{equation*}
$$

A minor simplification comes from the relation for the linear speed of sound in an ideal gas, $\gamma p_{o} / \rho_{0}=c^{2}$. This converts the preceding to

$$
\begin{equation*}
p=-\rho_{0}\left[\frac{\partial \phi}{\partial t}+\frac{1}{2} \nabla \phi \cdot \nabla \phi-\frac{1}{2 c^{2}}\left(\frac{\partial \phi}{\partial t}\right)^{2}\right] \tag{13.3.20}
\end{equation*}
$$

The absence of $\gamma$ or $\beta$ from this expression tells us that the nonlinear effect contained in this relation is kinematical in origin and may be used for both liquids and gases. We also see that the relation reduces to the linear one when the quadratic terms are dropped.

In summary, we have found that any acoustic signal in a liquid and all except the loudest acoustic signals in an ideal gas are governed by the quadratic version of the nonlinear wave equation, Eq. (13.3.18). Boundary conditions typically impose conditions on the particle velocity, which is $\bar{v}=\nabla \phi$. The pressure corresponding to a solution for $\phi$ is given by Eq. (13.3.20). In the exceptional case of a signal in an ideal gas whose amplitude is deemed to be too large to use a small-signal approximation, the appropriate nonlinear wave equation is Eq. (13.3.17) and Eq. (13.3.15) should be used to determine the pressure.

The derivation of the nonlinear wave equation is based on the assumptions that dissipation mechanisms are insignificant, and that the signal propagates isentropi-
cally. It is possible to modify this equation by adding terms that capture dissipative effects. Linear terms are adequate for this purpose, because dissipation, relaxation, and such are small effects that are comparable to the effect of nonlinearity. Suitable dissipation terms may be found in the modified versions of the linear wave equation that were developed in Sect.3.3.

Even if we exclude nonideal effects, a direct attempt to solve the nonlinear wave equation by standard techniques, such as separation of variables, would not be successful. For this reason, several simplified field equations have been developed. ${ }^{15}$ They differ in their assumptions regarding scales over which state variables change in various directions and the relative order of magnitude of terms representing various effects. Some model equations are particularly amenable to analysis by finite difference methods. Our concentration will be on systems that are amenable to analytical techniques. The quadratic nonlinear wave equation will be the framework for our investigations, but the same techniques may be applied to other model equations.

### 13.3.2 Frequency-Domain Formulation

Numerical techniques have been used to solve the nonlinear wave equation. The first thought might be to use finite difference formulas to approximate the derivatives in the nonlinear wave equation. However, that approach has inherent limitations, not the least of which is the accumulation of error with increasing distance and time. This is a crucial issue because the error might overwhelm the nonlinear effects, especially for a weakly nonlinear signal. Another issue is the large number of mesh points resulting from the requirements that the spatial mesh size be considerably smaller than a wavelength and the temporal time increment be a fraction of the interval over which the signal fluctuates.

The best computational approaches begin with an ansatz that captures as many aspects as possible of the fundamental physics of a system. If we are solely interested in situations where the excitation is periodic in time, a formulation that inherently recognizes the periodicity of the steady-state signal can reasonably be expected to be optimal. Periodicity allows the time signature to be represented by a Fourier series. In linear theory, each harmonic in that series would be a solution of the linear wave equation, and the series coefficients would be constants. We have seen with the Fubini-Ghiron solution that the amplitude of each harmonic in a plane wave is position dependent. This suggests that the periodic signal in any system may be described as a Fourier series whose harmonic functions are solutions of the linear wave equation for that system, and whose coefficients are position dependent. Making this ansatz satisfy the nonlinear wave equation will lead to a coupled set of first-order differential equations for the coefficients. Such equations are readily solved by standard numerical methods.

[^74]This approach is said to work in the frequency domain because the signal is decomposed into its harmonics. However, the most useful property of linear equations, which is the principle of superposition, is not applicable to nonlinear equations. A corollary is that one cannot simplify matters by using real or imaginary parts and factoring out complex exponential factors. We shall develop the basic approach by applying it to the analysis of a plane wave.

The period is $T$, so the fundamental frequency is $\omega_{1}=2 \pi / T$. The corresponding phase variable for a linear wave is

$$
\begin{equation*}
\theta_{1}=\omega_{1} t-k x, k=\frac{\omega_{1}}{c} \tag{13.3.21}
\end{equation*}
$$

Thus, the Fourier series representation of a plane wave is taken to be

$$
\begin{equation*}
\phi=\frac{1}{2} \sum_{n=-\infty}^{\infty} \Phi_{n}(x) e^{i n \theta_{1}}, \quad \Phi_{(-n)}=\Phi_{n}^{*} \tag{13.3.22}
\end{equation*}
$$

Equations for the $\Phi_{n}$ functions are obtained by requiring that this expression satisfies the quadratic nonlinear wave equation, Eq.(13.3.18). An important aspect of those operations arises when we consider $\nabla^{2} \phi$, which is

$$
\begin{equation*}
\nabla^{2} \phi=\frac{\partial^{2} \phi}{\partial x^{2}}=\frac{1}{2} \sum_{n=-\infty}^{\infty}\left[\Phi_{n}^{\prime \prime}-2 \operatorname{ink} \Phi_{n}^{\prime}-n^{2} k^{2} \Phi_{n}\right] e^{i n \theta_{1}} \tag{13.3.23}
\end{equation*}
$$

where a prime will denote an ordinary derivative with respect to $x$. Based on our experience thus far, as well as the expectation that nonlinear effects grow gradually, we consider $\left|\Phi_{n}^{\prime}(x)\right| \ll k\left|\Phi_{n}(x)\right|$. In other words, we take each of the amplitude factors to be a slowly varying function. Correspondingly, second derivatives of $\Phi_{n}$ are ignored, so the linear terms in the wave equation become

$$
\begin{equation*}
c^{2} \frac{\partial^{2} \phi}{\partial x^{2}}-\frac{\partial^{2} \phi}{\partial t^{2}}=-i c^{2} k \sum_{n=-\infty}^{\infty} n \Phi_{n}^{\prime} e^{i n \theta_{1}} \tag{13.3.24}
\end{equation*}
$$

The nonlinear terms in the wave equation, Eq. (13.3.18), require that derivatives of the Fourier series be multiplied. The manipulations are simplified by the slowness property, which allows derivatives of the $\Phi_{n}$ coefficients to be dropped when the product is formed. This means that only variation of the complex exponential is considered, so that

$$
\begin{equation*}
\frac{\partial \phi}{\partial t} \approx \frac{\partial \phi}{\partial \theta_{1}} \frac{\partial \theta_{1}}{\partial t}=\omega_{1} \frac{\partial \phi}{\partial \theta_{1}}, \frac{\partial \phi}{\partial x} \approx \frac{\partial \phi}{\partial \theta_{1}} \frac{\partial \theta_{1}}{\partial x}=-k \frac{\partial \phi}{\partial \theta_{1}} \tag{13.3.25}
\end{equation*}
$$

In turn, these properties lead to

$$
\begin{gather*}
\nabla^{2} \phi \frac{\partial \phi}{\partial t}=\frac{\partial^{2} \phi}{\partial x^{2}} \frac{\partial \phi}{\partial t} \approx k^{2} \omega_{1} \frac{\partial^{2} \phi}{\partial \theta_{1}^{2}} \frac{\partial \phi}{\partial \theta_{1}} \\
\frac{\partial}{\partial t}(\nabla \phi \cdot \nabla \phi)=2 \frac{\partial^{2} \phi}{\partial x \partial t} \frac{\partial \phi}{\partial x} \approx 2 k^{2} \omega_{1} \frac{\partial^{2} \phi}{\partial \theta_{1}^{2}} \frac{\partial \phi}{\partial \theta_{1}} \tag{13.3.26}
\end{gather*}
$$

The preceding relations allow us to combine the product terms in the wave equation. Each factor in the product must be described by a different summation index in order to assure that all individual products are described. Thus, the product terms become

$$
\begin{align*}
2(\beta-1) \nabla^{2} \phi \frac{\partial \phi}{\partial t}+\frac{\partial}{\partial t}(\nabla \phi \cdot \nabla \phi) & =2 \beta k^{2} \omega_{1} \frac{\partial^{2} \phi}{\partial \theta_{1}^{2}} \frac{\partial \phi}{\partial \theta_{1}} \\
& =-\frac{i}{2} \beta \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} m n^{2} k^{2} \omega_{1} \Phi_{m} \Phi_{n} e^{i(n+m) \theta_{1}} \tag{13.3.27}
\end{align*}
$$

Substitution of this expression and Eq. (13.3.24) into the nonlinear wave equation, Eq.(13.3.18), gives

$$
\begin{equation*}
-i c^{2} k \sum_{n=-\infty}^{\infty} n \Phi_{n}^{\prime} e^{i n \theta_{1}}=-\frac{i}{2} \beta k^{2} \omega_{1} \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} m n^{2} \Phi_{m} \Phi_{n} e^{i(n+m) \theta_{1}} \tag{13.3.28}
\end{equation*}
$$

The frequencies of the harmonic functions are multiples, so the functions are orthogonal over their common period. This property makes it possible to obtain a set of first-order differential equations for the amplitude functions. To that end, we multiply the equation by $\exp (-i j \theta)$, where $j$ may be any integer. We then integrate over a period, $-\pi<\theta_{1}<\pi$. All terms in the single summation, other than the one for which $n=j$, are orthogonal to this complex exponential. The only terms in the double summation that are not orthogonal to $\exp \left(-i j \theta_{1}\right)$ are those for which $m=j-n$. Thus, these operations filter a single term out of the single summation and reduce the double summation to a single one. The result is a set of coupled ordinary differential equations,

$$
\begin{equation*}
j k \Phi_{j}^{\prime}=\frac{1}{2} \beta \frac{k^{2} \omega}{c^{2}} \sum_{n=-\infty}^{\infty}(j-n) n^{2} \Phi_{(j-n)} \Phi_{n} \tag{13.3.29}
\end{equation*}
$$

It would be more meaningful if we worked with the Fourier pressure coefficients. We may describe $p(x, t)$ as a Fourier series,

$$
\begin{equation*}
p=\frac{1}{2} \sum_{n=-\infty}^{\infty} P_{n}(x) e^{i n \theta_{1}}, \quad P_{(-n)}=P_{n}^{*} \tag{13.3.30}
\end{equation*}
$$

The pressure is related to the velocity potential by $p=-\rho_{0} \partial \phi / \partial t$. (Nonlinearity is uniformly a second-order effect here.) The result of replacing $p$ and $\phi$ by their respective Fourier series and matching like harmonics shows that, for any $n$,

$$
\begin{equation*}
P_{n}=-i n \omega \rho_{0} \Phi_{n} \Longrightarrow n k \Phi_{n}=\left(\frac{i}{\rho_{0} c}\right) P_{n} \tag{13.3.31}
\end{equation*}
$$

When this relation is substituted into Eq. (13.3.29), the result is

$$
\begin{equation*}
\frac{P_{j}^{\prime}}{\rho_{0} c}=\frac{i}{2} \beta \frac{\omega}{c^{2}} \sum_{n=-\infty}^{\infty} n \frac{P_{(j-n)} P_{n}}{\left(\rho_{0} c\right)^{2}} \tag{13.3.32}
\end{equation*}
$$

This form suggests nondimensionalization of the position and pressure coefficients according to

$$
\begin{equation*}
\tilde{x}=k x, \quad \tilde{P}_{n}=\frac{P_{n}}{\rho_{0} c^{2}} \tag{13.3.33}
\end{equation*}
$$

This replacement converts the differential equations to

$$
\begin{equation*}
\frac{d \tilde{P}_{j}}{d \tilde{x}}=\frac{i}{2} \beta \sum_{n=-\infty}^{\infty} n \tilde{P}_{(j-n)} \tilde{P}_{n} \tag{13.3.34}
\end{equation*}
$$

These differential equations apply for all $j$, but it would be repetitive to evaluate the coefficients for $j<0$ because $\tilde{P}_{(-j)}$ is the complex conjugate of $\tilde{P}_{j}$. If we are to make use of this property to reduce the number of variables, we must explicitly account for the complex conjugate property within the summation. The first step is to break up the summation into two parts, $n>0$ and $n<0$. (The contribution from $n=0$ is zero.) For the second part, $n$ is replaced by $-n$, which gives

$$
\begin{equation*}
\frac{d \tilde{P}_{j}}{d \tilde{x}}=\frac{i}{2} \beta\left[\sum_{n=1}^{\infty} n \tilde{P}_{(j-n)} \tilde{P}_{n}-\sum_{n=1}^{\infty} n \tilde{P}_{(j+n)} \tilde{P}_{(-n)}\right] \tag{13.3.35}
\end{equation*}
$$

In the equation for $j=0$, the sums cancel, so that

$$
\begin{equation*}
\frac{d \tilde{P}_{0}}{d \tilde{x}}=0 \tag{13.3.36}
\end{equation*}
$$

In other words, the mean pressure $\tilde{P}_{0}$ is independent of the position.
For $j>0$, the second summation in Eq.(13.3.35) features a coefficient with a negative subscript. The complex conjugate property gives $\tilde{P}_{(-n)}=\tilde{P}_{n}^{*}$. The first summation also has a coefficient with a negative subscript when $n>j$. This is handled by splitting the summation into two ranges: $n \leq j$ and $n>j$, with $\tilde{P}_{(j-n)}=\tilde{P}_{(n-j)}^{*}$ for the latter range. These operations lead to

$$
\begin{equation*}
\frac{d \tilde{P}_{j}}{d \tilde{x}}=\frac{i}{2} \beta\left[\sum_{n=1}^{j} n \tilde{P}_{(j-n)} \tilde{P}_{n}+\sum_{n=j+1}^{\infty} n \tilde{P}_{(n-j)}^{*} \tilde{P}_{n}-\sum_{n=1}^{\infty} n \tilde{P}_{(j+n)} \tilde{P}_{n}^{*}\right], j>0 \tag{13.3.37}
\end{equation*}
$$

Replacing $n$ with $j+n$ in the middle summation reduces this expression to

$$
\begin{equation*}
\frac{d \tilde{P}_{j}}{d \tilde{x}}=\frac{i}{2} \beta\left[\sum_{n=1}^{j} n \tilde{P}_{(j-n)} \tilde{P}_{n}+\sum_{n=1}^{\infty} j \tilde{P}_{n}^{*} \tilde{P}_{(j+n)}\right], j>0 \tag{13.3.38}
\end{equation*}
$$

This is the form we seek because it only features $P_{n}$ values for $n>0$.
The first summation in this equation describes the growth or decay of harmonic $j$ due to the interaction of harmonics below $j$. The second summation captures the influence on harmonic $j$ of a higher harmonic interacting with all harmonics. For example, the first summation indicates that second harmonics $(j=2)$ are generated by the interaction of the first harmonic with itself, while the second summation indicates that second harmonics are also generated by the interaction of many harmonic pairs consisting of first and third, second and fourth, and so on.

The differential equations must be solved numerically, so it is necessary to truncate the pressure series. Let $N$ be the highest harmonic that is retained. Thus, we set $\tilde{P}_{n}=0$ for $n>N$. Neither subscript in the first summation will exceed $N$, so that summation requires no adjustment. The lowest value of $n$ in the second summation that leads to a factor being zero is that for which $j+n=N+1$. Thus, that summation is truncated at $n=N-j$, which means that the equations to be solved numerically are

$$
\begin{equation*}
\frac{d \tilde{P}_{j}}{d \tilde{x}}=\frac{i}{2} \beta\left[\sum_{n=1}^{j} n \tilde{P}_{j-n} \tilde{P}_{n}+\sum_{n=1}^{N-j} j \tilde{P}_{n}^{*} \tilde{P}_{(j+n)}\right], \quad j=1,2, \ldots, N \tag{13.3.39}
\end{equation*}
$$

The second summation is not executed if $j=N$. The value of $\tilde{P}_{0}$ is set at whatever constant value it has at $x=0$, so there are $N$ pressure coefficients to determine and $N$ differential equations to solve.

An attractive feature of the frequency-domain analysis is its generality. Monitoring the term on the right side of Eq. (13.3.39) enables one to delve into the interaction of harmonics at a fundamental level, rather than merely as a computed result. Another beneficial aspect of the formulation is that it is readily modified to incorporate the effects of dissipation. Section 3.3 discusses how dissipation is manifested in the frequency domain. It was shown that weak dissipation attenuates the complex ampli-
tude of a plane wave. Dissipation is important for nonlinear propagation because the decay constant typically increases with increasing frequency. Thus, attenuation due to dissipation counters the nonlinear tendency for harmonics to grow as the wave propagates.

For propagation in the direction of increasing $x$, the attenuation factor is $e^{-\alpha x}$. The absorption coefficient $\alpha$ is a real quantity that depends on frequency. Different dissipation effects, such as bulk viscosity or wall friction, lead to different dependencies, so we will write it as $\alpha(\omega)$. If nonlinear effects were not present, this exponential decay would be obtained if the coefficients were governed by

$$
\begin{equation*}
\frac{d \tilde{P}_{j}}{d(\tilde{x} / k)}+\alpha\left(j \omega_{1}\right) \tilde{P}_{j}=0 \tag{13.3.40}
\end{equation*}
$$

The assumption that dissipation is weak implies that it is not altered by nonlinear effects, and vice versa. Consequently, we may insert the dissipation term into the nonlinear amplitude equations, with the result that Eq.(13.3.39) becomes

$$
\begin{equation*}
\frac{d \tilde{P}_{j}}{d \tilde{x}}=-\frac{\alpha\left(j \omega_{1}\right)}{k} \tilde{P}_{j}+\frac{i}{2} \beta\left[\sum_{n=1}^{j} n \tilde{P}_{j-n} \tilde{P}_{n}+\sum_{n=1}^{N-j} j \tilde{P}_{n}^{*} \tilde{P}_{(j+n)}\right], j=1,2, \ldots, N \tag{13.3.41}
\end{equation*}
$$

Note that the presence of dissipation does not alter the fact that $\tilde{P}_{0}$ is independent of $x$.

Starting values for these differential equations are obtained from the boundary condition. In the weakly nonlinear regime, $\varepsilon \ll 1$, the particle velocity at $x=0$ equals the motion of the boundary. Let $\varepsilon c V(t)$ denote the velocity there. The smallsignal approximation gives the corresponding pressure $\rho_{0} c(\varepsilon c V)$. The phase variable at $x=0$ is $\theta=\omega_{1} t$, so matching the Fourier series to the motion that is imposed gives

$$
\begin{equation*}
\frac{1}{2} \rho_{0} c^{2} \sum_{n=-\infty}^{\infty} \tilde{P}_{n}(\tilde{x}=0) e^{i n \omega_{1} t}=\rho_{0} c[\varepsilon c V(t)] \tag{13.3.42}
\end{equation*}
$$

Orthogonality of the harmonic functions in a Fourier series then leads to

$$
\begin{equation*}
\tilde{P}_{j}(x=0)=\frac{\varepsilon}{2 \pi} \int_{-\pi / \omega_{1}}^{\pi / \omega_{1}} V(t) e^{-i j \omega_{1} t} d\left(\omega_{1} t\right) \tag{13.3.43}
\end{equation*}
$$

EXAMPLE 13.5 Determine the Fourier series pressure coefficients for the steady-state plane wave generated by a harmonic source at $x=0$. The particle velocity at that location is $v=\varepsilon c \sin (\omega t)$ with $\omega=2 \pi(5000) \mathrm{rad} / \mathrm{s}$ and
$\varepsilon=2\left(10^{-5}\right)$. The fluid is air at standard conditions. Let the decay constant be $\alpha=2.54\left(10^{-13}\right) \omega^{2}$ neper $/ \mathrm{m}$. Carry out the evaluation to the distance where a shock would form if there were no dissipations, which is $\left(x_{\text {shock }}\right)_{\text {sine }}=1 /(\beta \varepsilon k)$. Compare the results to those obtained from the FubiniGhiron solution for the dissipationless case. Also evaluate the waveform at $\left(x_{\text {shock }}\right)_{\text {sine }}$.

## Significance

Issues such as implementation of a numerical procedure for solving the coupled differential equations and the number of harmonics that are required shall be covered. Equally important is the demonstration of the interplay of dissipation and nonlinearity effects.

## Solution

The analysis entails solving Eq. (13.3.41) subject to the initial conditions described by Eq. (13.3.43). The velocity at $x=0$ is specified to be $v=\varepsilon c\left(e^{i \omega t}-e^{-i \omega t}\right) /(2 i)$. Matching $\rho_{0} c v$ to the Fourier series for $p$ at $x=0$ leads to

$$
\begin{equation*}
\tilde{P}_{1}(x=0)=-i \varepsilon, \text { otherwise } P_{n}(x=0)=0 \tag{1}
\end{equation*}
$$

The mean value, $\tilde{P}_{0}$, is zero. Nevertheless, we shall treat it as a variable because doing so leads to the implementation of a numerical procedure that is valid for any (periodic) boundary velocity.

The coupled differential equations may be solved with a Runge-Kutta numerical integrator. Several versions are available; we shall use a forth-fifth order scheme. The form that is required is

$$
\begin{equation*}
\frac{d \tilde{P}_{j}}{d \tilde{x}}=F_{j}\left(\tilde{P}_{n}, \tilde{x}\right), j=0, \ldots, N \tag{2}
\end{equation*}
$$

Equation (13.3.41) matches this general form, with
$F_{0}\left(\tilde{P}_{n}, x\right)=0$
$F_{j}\left(\tilde{P}_{n}, x\right)=-\frac{\alpha\left(j \omega_{1}\right)}{k} \tilde{P}_{j}+\frac{i}{2} \beta\left[\sum_{n=1}^{j} n \tilde{P}_{j-n} \tilde{P}_{n}+\sum_{n=1}^{N-j} j \tilde{P}_{n}^{*} \tilde{P}_{(j+n)}\right], j=1, \ldots, N-1$
$F_{j}\left(\tilde{P}_{n}, x\right)=-\frac{\alpha\left(N \omega_{1}\right)}{k} \tilde{P}_{j}+\frac{i}{2} \beta\left[\sum_{n=1}^{N} n \tilde{P}_{j-n} \tilde{P}_{n}\right], j=N$

There is no simple way to vectorize computation of the $F_{j}$ terms, so the operations are implemented term by term. Within a loop that increments $j$ from 1 to $N$, the computation begins by equating $F_{j}$ to the linear term. A loop over $n$ from 1 to $j$ adds the terms in the first summation. Then, if $j$ is less than $N$, a loop over $n$ from one to $N-j$ adds the terms in the summations. The only caution is that if one's software requires that subscripts begin at one, as is the case with MATLAB, then the subscript number must be one greater than the associated harmonic number.

The dissipation factor that appears in the differential equations is $\alpha\left(j \omega_{1}\right) k$. For the given $\alpha$ dependence, this is

$$
\begin{equation*}
\frac{\alpha(j \omega)}{k}=2.54\left(10^{-13}\right) \frac{j^{2} \omega_{1}^{2}}{k}=2.54\left(10^{-13}\right)\left(\omega_{1} c\right) j^{2} \tag{4}
\end{equation*}
$$

The units of $\alpha$ are nepers $/ \mathrm{m}$, so $\alpha / k$ is a dimensionless quantity. This means that the numerical coefficient has units of $\mathrm{s}^{2} / \mathrm{m}$.

The computation requires selection of $N$. We know from Fig. 13.8 that $\left|\tilde{P}_{n}\right| \ll\left|\tilde{P}_{1}\right|$ if $n>5$, and dissipation will affect the higher harmonics more than $\tilde{P}_{1}$ because the dissipation factor increases monotonically with increasing harmonic number. We shall begin with $N=10$, which would seem to be adequate to assure that all significant nonlinear interactions are properly described. As one should always do when making an a priori decision like this, we will verify the adequacy of this truncation by comparing the results to those obtained from a larger $N$.

It is requested that the series coefficients be compared to those of the FubiniGhiron solution of the dissipationless case. That solution is described in Eq. (13.2.13) in terms of the retarded time $t^{\prime}=t-x / c$ and $x$ nondimensionalized relative to the shock formation distance for harmonic excitation, $\left(x_{\text {shock }}\right)_{\text {sine }}=1 /(\beta \varepsilon k)$. In terms of the present variables, the complex Fourier series for the signal in the case of an ideal fluid is

$$
\begin{equation*}
p=\frac{1}{2} \rho_{0} c^{2} \sum_{n=1}^{\infty} \tilde{P}_{n} e^{i n \theta_{1}}, \quad \tilde{P}_{n}=\frac{2 J_{n}(n \beta \varepsilon \tilde{x})}{\operatorname{in\beta } \tilde{x}} \frac{P_{n}}{\rho_{0} c^{2}}=\frac{2 \varepsilon}{n \sigma} J_{n}(n \sigma) \tag{5}
\end{equation*}
$$

One computational check is to compare the solution of the differential equations for $\alpha=0$ to this expression. Evaluation of $\left|\left(\tilde{P}_{n}\right)_{\text {comp }}-\left(\tilde{P}_{n}\right)_{\mathrm{F}-\mathrm{G}}\right| / \varepsilon$ showed an error that rose from $0.02 \%$ for $n=1$ to $0.14 \%$ for $n=9$. The error for $n=N=10$ was found to be $4.8 \%$, which is typical behavior when a Fourier series is truncated. For reasons that will become apparent when we examine waveforms, the computation was repeated with $N=50$. The error for the first ten harmonics in that case was found to be less than $0.005 \%$.

The pressure coefficients when $\alpha(j \omega)$ is given by Eq. (4) are the solid curves in Fig. 1. The dashed curves are the values for the dissipationless case. Each amplitude is attenuated. For clarity, both linear and logarithmic scales are used. Viewed as a ratio, the loss is greater as the harmonic number increases. Whether this is a drastic effect depends on the appearance of the waveform at the shock formation distance.


Figure 1.

A waveform is obtained by evaluating Eq.(13.3.30). The result at the shock formation distance, which is $\tilde{x}=k\left(x_{\text {shock }}\right)_{\text {sine }}=1 /(\beta \varepsilon)$, appears in Fig. 2. The result for the ideal case could be obtained by evaluating the Riemann solution; instead, it also was obtained by computing the series. The cutoff number $N$ was raised to obtain these waveforms because a much higher resolution is required to capture the steep slope of the $\alpha=0$ curve. This is unlike the behavior of the curve for $\alpha \neq 0$, which was found to differ negligibly from the one obtained from $N=10$.


Figure 2.

The waveform in Fig. 2 shows that the presence of dissipation has two effects. The first is to attenuate the overall amplitude, which would be $\left|p / \rho_{0} c^{2}\right|=\varepsilon$ if $\alpha$ were zero. The second effect is to reduce the slope in the phase where $\partial p / \partial t>0$. Thus, it forestalls the formation of a shock. The differential equations lose validity when a shock forms because they do not account for the Rankine-Hugoniot conditions. Conversely, they can be used for any $\tilde{x}$ in the $\alpha \neq 0$ case, provided $\partial p / \partial t$ remains finite at lesser $\tilde{x}$. Figure 3 is the waveform at $\tilde{x}=4 /(\beta \varepsilon)$. It appears that the attenuation is sufficiently great that a shock will not form at any location. Because the attenuation increases with increasing harmonic number, at a sufficiently large distance the signal will be reduced to a greatly attenuated fundamental. This is one of the reasons that sounds that are generated in everyday life do not form shocks.


Figure 3.

### 13.3.3 Regular Perturbation Series Expansion

There are a number of fundamental systems and phenomena that can be solved analytical, rather than by recourse to numerical methods. One approach that has had success is a perturbation series representation of the pressure or velocity potential. The concept is limited by the requirement that the linearized version of the problem be solvable. In actuality, this is not much of a limitation, because it is unreasonable to expect to solve the nonlinear equations if the linearized version of those equations cannot be solved. The starting point is identification of a basic parameter that can be used as a metric of the relative importance of linear and nonlinear terms in the field equation. We have already encountered such a parameter, the acoustic Mach number $\varepsilon$.

In a linear problem, any constant that multiplies the excitation scales the response in the same way. This is not true for a nonlinear signal, so it follows that the potential function and therefore, the acoustic pressure and particle velocity are unknown functions of the Mach number $\varepsilon$, as well as position and time. This dependence on $\varepsilon$ can be expected to be analytical, at least in regions where shocks do no occur. An analytical dependence on $\varepsilon$ implies that the variables can be represented as a Taylor series relative to $\varepsilon=0$. The response variable for the nonlinear wave equation is $\phi$, so its Taylor series is

$$
\begin{equation*}
\phi(\bar{x}, t, \varepsilon)=\left.\phi(\bar{x}, t, \varepsilon)\right|_{\varepsilon=0}+\left.\varepsilon\left(\frac{\partial}{\partial \varepsilon} \phi(\bar{x}, t, \varepsilon)\right)\right|_{\varepsilon=0}+\left.\frac{1}{2} \varepsilon^{2}\left(\frac{\partial^{2}}{\partial \varepsilon^{2}} \phi(\bar{x}, t, \varepsilon)\right)\right|_{\varepsilon=0}+\cdots \tag{13.3.44}
\end{equation*}
$$

The functional dependence of $\phi$ is not known when the analysis is initiated, so the derivatives may be replaced by functions to be determined. Furthermore, because $\varepsilon$ is the scaling factor of the excitation, setting $\varepsilon=0$ means there is no excitation, in which case $\phi$ will be zero. This means that the series should not contain a term that is independent of $\varepsilon$. Thus, the series reduces to

$$
\begin{equation*}
\phi(\bar{x}, t, \varepsilon)=\varepsilon \phi_{1}(\bar{x}, t)+\varepsilon^{2} \phi_{2}(\bar{x}, t)+\varepsilon^{3} \phi_{3}(\bar{x}, t)+\cdots \tag{13.3.45}
\end{equation*}
$$

When this series is used as a Taylor series, the concept is to add terms until it converges. Adding more terms will decrease the error of the truncated series relative to the true $\phi$. This is not the viewpoint if the representation is considered to be a perturbation series. Rather, the number of terms is fixed. If $|\varepsilon| \ll 1$, we know that the difference between the series that is truncated beyond $\varepsilon^{N}$ and the true $\phi$ will be proportional to $\varepsilon^{N+1}$. In other words, lengthening a perturbation solution would reduce the error at a specified $\varepsilon$, rather than increasing the range of $\varepsilon$ for which the error is less than a specified amount. The change in viewpoint from a Taylor series to a perturbation series is fortunate, because the analysis of each function $\phi_{n}$ becomes progressively more difficult as $n$ increases. It follows that there is no intent to increase the length of the series, even though doing so might lead to a convergent answer.

Frequent reference will be made to the order of magnitude of a term, so it is important to have a clear definition of that metric. A function $f(\varepsilon)$ is said to be $O\left(\varepsilon^{n}\right)$ if, for two very small $\varepsilon$ values, it is found that $f\left(\varepsilon_{1}\right) / f\left(\varepsilon_{2}\right)$ is essentially $\left(\varepsilon_{1} / \varepsilon_{2}\right)^{n}$. We will use the weakly nonlinear version of the wave equation, Eq. (13.3.18). Cubic terms have been omitted from this equation. The velocity potential is $O(\varepsilon)$, so the omitted terms are $O\left(\varepsilon^{3}\right)$. It would be inconsistent to seek terms at that order in the response, so the perturbation series for the velocity potential pressure and particle velocity are truncated at $O\left(\varepsilon^{2}\right)$. The forms are

$$
\begin{align*}
& \phi=\varepsilon \phi_{1}(x, t)+\varepsilon^{2} \phi_{2}(x, t)+\cdots \\
& p=\varepsilon p_{1}(x, t)+\varepsilon^{2} p_{2}(x, t)+\cdots \tag{13.3.46}
\end{align*}
$$

We have no intrinsic interest in the density as a response variable, but if we did, its perturbation series would begin with $\rho_{0}$, which is the $O(1)$ term.

The next step is to assemble the equations to be satisfied. For a planar wave, they are

$$
\begin{gather*}
v=\frac{\partial \phi}{\partial x} \\
p=-\rho_{0}\left[\frac{\partial \phi}{\partial t}+\frac{1}{2}\left(\frac{\partial \phi}{\partial x}\right)^{2}-\frac{1}{2 c^{2}}\left(\frac{\partial \phi}{\partial t}\right)^{2}\right]  \tag{13.3.47}\\
c^{2} \frac{\partial^{2} \phi}{\partial x^{2}}-\frac{\partial^{2} \phi}{\partial t^{2}}-\left[2(\beta-1) \frac{\partial^{2} \phi}{\partial x^{2}} \frac{\partial \phi}{\partial t}+2 \frac{\partial \phi}{\partial x} \frac{\partial^{2} \phi}{\partial x \partial t}\right]=0
\end{gather*}
$$

The early development of the method of characteristics showed that the effect of a moving boundary condition is always $O\left(\varepsilon^{2}\right)$, see Eq. (13.1.17). Nevertheless, it is useful to see how such effects may be incorporated in a perturbation analysis. We have defined $\varepsilon$ to measure the magnitude of the boundary motion, so we take the displacement of the boundary to be $\varepsilon X(t)$. Thus, the boundary condition is

$$
\begin{equation*}
\left.\frac{\partial \phi}{\partial x}\right|_{x=\varepsilon X(t)}=\varepsilon \dot{X}(t) \tag{13.3.48}
\end{equation*}
$$

Statement of the problem is completed by imposing the radiation condition, which requires that $p$ be an outgoing wave as $x \rightarrow \infty$, and that the fluid was at the ambient state for $t<0$, so that $p=v=0$ at $t=0$.

The fundamental step in the analysis stems from the requirement that the perturbation series constitute a solution for any $\varepsilon$. When any such series is substituted into an equation, terms like order of $\varepsilon$ may be collected. The result will have the form $\varepsilon \times($ term 1$)+\varepsilon^{2} \times(\operatorname{term} 2)=0$. It is necessary that this equation be satisfied for many different values of $\varepsilon$. This requirement can be met only if the coefficient of each power of $\varepsilon$ is zero. Thus, a single equation becomes an equation for each power of $\varepsilon$ : term $1=0$, term $2=0, \ldots$. An equivalent statement is that the coefficients of like powers of $\varepsilon$ in each equation must match.

In this way, we obtain a sequence of problems in which $\varepsilon$ does not appear. Each set of equations is described as being the equations at the order of the associated power of $\varepsilon$. The first-order set consists of term 1 from each governing equation, the second-order set consists of term 2 for each equation, and so on. Thus, the first-order equations derived from Eq. (13.3.47) are

$$
\begin{gather*}
v_{1}=\frac{\partial \phi_{1}}{\partial x}  \tag{13.3.49}\\
p_{1}=-\rho_{0} \frac{\partial \phi_{1}}{\partial t} \\
c^{2} \frac{\partial^{2} \phi_{1}}{\partial x^{2}}-\frac{\partial^{2} \phi_{1}}{\partial t^{2}}=0
\end{gather*}
$$

The second-order equations are

$$
\begin{gather*}
v_{2}=\frac{\partial \phi_{2}}{\partial x}  \tag{13.3.50}\\
p_{2}=-\rho_{0} \frac{\partial \phi_{2}}{\partial t}-\left[\frac{1}{2}\left(\frac{\partial \phi_{1}}{\partial x}\right)^{2}-\frac{1}{2 c^{2}}\left(\frac{\partial \phi_{1}}{\partial t}\right)^{2}\right] \\
c^{2} \frac{\partial^{2} \phi_{2}}{\partial x^{2}}-\frac{\partial^{2} \phi_{2}}{\partial t^{2}}=2(\beta-1) \frac{\partial^{2} \phi_{1}}{\partial x^{2}} \frac{\partial \phi_{1}}{\partial t}+2 \frac{\partial \phi_{1}}{\partial x} \frac{\partial^{2} \phi_{1}}{\partial x \partial t}
\end{gather*}
$$

A slightly simpler differential equation for $\phi_{2}$ is obtained if the first-order equation is used to replace $\partial^{2} \phi_{1} / \partial x^{2}$ with $\left(1 / c^{2}\right) \partial^{2} \phi_{1} / \partial t^{2}$. Doing so leads to

$$
\begin{equation*}
c^{2} \frac{\partial^{2} \phi_{2}}{\partial x^{2}}-\frac{\partial^{2} \phi_{2}}{\partial t^{2}}=\frac{\partial}{\partial t}\left[\frac{(\beta-1)}{c^{2}}\left(\frac{\partial \phi_{1}}{\partial t}\right)^{2}+\left(\frac{\partial \phi_{1}}{\partial x}\right)^{2}\right] \tag{13.3.51}
\end{equation*}
$$

It will be noted that at each order the terms that are lower order have been placed to the right of the equality sign. Doing so anticipates the way in which the equations are
solved. The first-order equations are linear in the dependent variables, which means we should be able to solve them. The left side of the first- and second-order equations is the same, but each second-order equation is inhomogeneous. Substitution of the first-order solutions will convert them into terms that represent a spatial distribution of sources.

The moving boundary condition also is transformed to a set of perturbation equations by means of a Taylor series relative to $x=0$, specifically,

$$
\begin{equation*}
\left.\frac{\partial \phi}{\partial x}\right|_{x=\varepsilon X(t)}=\left.\frac{\partial \phi}{\partial x}\right|_{x=0}+\left.\varepsilon X(t) \frac{\partial^{2} \phi}{\partial x^{2}}\right|_{x=0}+\left.\frac{1}{2}(\varepsilon X(t))^{2} \frac{\partial^{2} \phi}{\partial x^{2}}\right|_{x=0}+\cdots=\varepsilon \dot{X}(t) \tag{13.3.52}
\end{equation*}
$$

Only effects up to $O\left(\varepsilon^{2}\right)$ are of interest. Thus, we substitute the perturbation series for $\phi$ and match the $O(\varepsilon)$ and $O\left(\varepsilon^{2}\right)$ terms on either side of the equality. The result is a sequence of boundary conditions,

$$
\begin{gather*}
\left.\frac{\partial \phi_{1}}{\partial x}\right|_{x=0}=\dot{X}(t)  \tag{13.3.53}\\
\left.\frac{\partial \phi_{2}}{\partial x}\right|_{x=0}=-\left.X(t) \frac{\partial^{2} \phi_{1}}{\partial x^{2}}\right|_{x=0}
\end{gather*}
$$

In addition, the radiation and initial conditions must be satisfied for any $\varepsilon$, which leads to the requirements that each order be an outgoing wave, and that each order of $p$ and $v$ be zero for $t<0$.

In most systems, solution of the linear problem requires specification of the excitation. However, the d'Alembert solution, which is general, is available for plane waves. An outgoing wave that satisfies the linear wave equation in Eq.(13.3.49) is $\phi_{1}=f(t-x / c)$. The function $f$ is determined by satisfying the first-order boundary condition, Eq. (13.3.53),

$$
\begin{equation*}
-\frac{1}{c} f^{\prime}(t)=\dot{X} \tag{13.3.54}
\end{equation*}
$$

The notation $f^{\prime}$ prime denotes differentiation with respect to the argument of $f$, which is $t-x / c$. The chain rule for differentiation leads to $f^{\prime}(t-x / c) \equiv(\partial / \partial t) f(t-x / c)$, so the boundary condition may be integrated with respect to time to find that

$$
\begin{equation*}
\phi_{1}=-c X\left(t-\frac{x}{c}\right) \tag{13.3.55}
\end{equation*}
$$

The corresponding state variables are

$$
\begin{equation*}
v_{1}=\dot{X}\left(t-\frac{x}{c}\right), \quad p_{1}=\rho_{0} c \dot{X}\left(t-\frac{x}{c}\right) \tag{13.3.56}
\end{equation*}
$$

It is implicit to this solution that $X(t)=0$ for $t<0$, so these terms are consistent with the initial condition.

Now that we have determined the first-order terms, we proceed to the next order. Substitution of $\phi_{1}$ into the second-order wave equation, Eq. (13.3.51), gives

$$
\begin{equation*}
c^{2} \frac{\partial^{2} \phi_{2}}{\partial x^{2}}-\frac{\partial^{2} \phi_{2}}{\partial t^{2}}=\beta \frac{\partial}{\partial t}\left[\dot{X}\left(t-\frac{x}{c}\right)^{2}\right] \tag{13.3.57}
\end{equation*}
$$

The corresponding second-order boundary condition is

$$
\begin{equation*}
\left.\frac{\partial \phi_{2}}{\partial x}\right|_{x=0}=\frac{1}{c} X(t) \ddot{X}(t) \tag{13.3.58}
\end{equation*}
$$

The solution of the second-order wave equation consists of a complementary equation solution $\left(\phi_{2}\right)_{\mathrm{c}}$ and a particular solution $\left(\phi_{2}\right)_{\mathrm{p}}$. The definition of $\left(\phi_{2}\right)_{\mathrm{p}}$ is that its substitution into the left side of the equation exactly matches the right side, whereas substitution of $\left(\phi_{2}\right)_{c}$ gives zero. The left side of the equation is the linear wave equation, so any term that gives zero on the right side must be a function of either characteristic variable. The requirement that such a term be a forward propagating wave removes dependence on the backward characteristic. Therefore, we know that

$$
\begin{equation*}
\left(\phi_{2}\right)_{\mathrm{c}}=F_{\mathrm{c}}\left(t-\frac{x}{c}\right) \tag{13.3.59}
\end{equation*}
$$

We now are presented with a small dilemma, because the source term on the right side of Eq. (13.3.57) is a function of $t-x / c$. Any function with such dependence is a homogeneous solution, so we cannot use that combination of variables to form a particular solution. Let us try a particular solution that is modified by introducing a factor $x$, specifically,

$$
\begin{equation*}
\left(\phi_{2}\right)_{\mathrm{p}}=x F_{\mathrm{p}}\left(t-\frac{x}{c}\right) \tag{13.3.60}
\end{equation*}
$$

Substitution of this ansatz into the second-order wave equation yields

$$
\begin{equation*}
-2 c F_{\mathrm{p}}^{\prime}\left(t-\frac{x}{c}\right)=\beta \frac{\partial}{\partial t}\left[\dot{X}\left(t-\frac{x}{c}\right)^{2}\right] \tag{13.3.61}
\end{equation*}
$$

According to the chain rule for differentiation, $(\partial / \partial t) F_{\mathrm{p}}(t-x / c)=F_{\mathrm{p}}^{\prime}(t-x / c)$. Thus, the preceding may be integrated directly. The resulting particular solution is

$$
\begin{equation*}
\left(\phi_{2}\right)_{\mathrm{p}}=-\frac{\beta}{2 c} x \dot{X}\left(t-\frac{x}{c}\right)^{2} \tag{13.3.62}
\end{equation*}
$$

One might wonder at this juncture why the trial form of the particular solution was not taken to be $t F_{\mathrm{p}}(t-x / c)$, which could also be made to match the source term. The term that was selected could be justified by noting that it matches the behavior observed in the Riemann solution, specifically that distortion of the pressure profile increases with increasing $x$. However, invocation of this argument would violate
the intention of solving the problem without having prior knowledge of the Riemann solution. A justification fitting this criterion is that the form selected for the particular solution must be valid for any $X(t)$. If $X(t)$ is periodic for $t>0$, then regardless of the fact that the system is nonlinear, the response must evolve to have that same temporal periodicity. Such a property cannot be obtained if $t$ is a factor in the trial solution.

The function $F_{\mathrm{c}}$ for the complementary solution in Eq.(13.3.59) is arbitrary in regard to satisfying the second-order wave equation. It is set by satisfying the secondorder boundary condition, Eq. (13.3.58). This condition must be satisfied by the sum of the complementary and particular solutions, which is why the particular solution for $\phi_{2}$ was found first. Thus, we seek a function $F_{\mathrm{c}}(t-x / c)$ for which

$$
\begin{equation*}
\left.\frac{\partial}{\partial x}\left[\left(\phi_{2}\right)_{\mathrm{c}}+\left(\phi_{2}\right)_{\mathrm{p}}\right]\right|_{x=0} \equiv-\frac{1}{c} F_{\mathrm{c}}^{\prime}(t)-\frac{\beta}{2 c} \dot{X}(t)^{2}=\frac{1}{c} X(t) \ddot{X}(t) \tag{13.3.63}
\end{equation*}
$$

Because $F_{\mathrm{c}}^{\prime}(t) \equiv \dot{F}_{\mathrm{c}}(t)$ at $x=0$, this equation may be integrated, although the integral can be evaluated only upon specification of $X(t)$. Thus, we find that

$$
\begin{equation*}
F_{\mathrm{c}}(t)=-\int_{0}^{t}\left[\frac{\beta}{2} \dot{X}\left(t^{\prime}\right)^{2} d t^{\prime}+X\left(t^{\prime}\right) \ddot{X}\left(t^{\prime}\right)\right] d t^{\prime} \tag{13.3.64}
\end{equation*}
$$

The second-order potential at any $x$ is obtained by replacing $t$ with $t-x / c$ in this expression and then combining it with the particular solution. Doing so yields

$$
\begin{equation*}
\phi_{2}=-\frac{\beta}{2 c} x \dot{X}\left(t-\frac{x}{c}\right)^{2}-\int_{0}^{t-x / c}\left[\frac{\beta}{2} \dot{X}\left(t^{\prime}\right)^{2} d t^{\prime}+X\left(t^{\prime}\right) \ddot{X}\left(t^{\prime}\right)\right] d t^{\prime} \tag{13.3.65}
\end{equation*}
$$

The second-order pressure and particle velocity may be obtained from Eq. (13.3.50). The result shall not be listed here because doing so would not advance the dual objective of demonstrating how to implement a perturbation analysis, and how to handle a moving boundary condition. However, the expression for $\phi_{2}$ that we have derived serves several purposes.

It is evident that the full solution, which is obtained by combining first- and secondorder effects, is not at all like the Riemann solution. The combination is not consistent with the small-signal approximation that $p=\rho_{0} c v$. However, the discrepancy is $O\left(\varepsilon^{2}\right)$ everywhere, so the difference is not significant. It is far more important that the perturbation solution has a limited range of validity. This is so because $\left(\phi_{2}\right)_{\mathrm{p}}$ is proportional to $x$, so it grows in magnitude. In the vicinity of the shock formation distance, $x=O(1 / \varepsilon)$. This makes $\varepsilon^{2} \phi_{2}$ have the same magnitude as $\varepsilon \phi_{1}$, but the basic concept of a perturbation series is that each succeeding term is small compared to the preceding terms. Therefore, the velocity potential we have derived is only valid in regions where $x=O(1)$. A perturbation series that is not equally accurate at all locations is said to be nonuniformly valid. In contrast, the Riemann solution is valid everywhere up to the shock formation distance. Some individuals say that $\phi_{2}$ has a singularity at $x=\infty$ because it grows without bound as $x$ increases.

Despite these faults, a nonuniformly valid series might be quite useful. Suppose we were interested in mapping the field around the boundary and we had no knowledge of the Riemann solution. The perturbation solution would explain certain nonlinear effects, such as the occurrence of second harmonics if $X(t)$ is sinusoidal, and the fact that the difference between the linear and nonlinear solutions grows with increasing propagation distance. The nonuniformly valid solution also illustrates a general fact that nonlinear effects arising from moving boundary conditions do not grow with increasing distance from the boundary. Ultimately, the most important justification for the analysis is that its faults can be corrected. Methods for doing so are the subject of the next section.

EXAMPLE 13.6 A plane wave is induced by a wall vibration that constitutes a beat. The steady state motion is $v=\varepsilon c \sin (\sigma t) \sin (\omega t)$, where $\sigma \ll \omega$. Identify the terms in the second-order velocity potential that are not uniformly valid.

## Solution

An analysis based on the d'Alembert solution, like the one for the preceding development, will not be suitable for more general systems. Frequency-domain tools often are likely to be employed in such circumstances. We will see how to do so here.

## Solution

Although the basic perturbation equations have already been stated, it is good practice to assemble them anew. We seek a solution for the velocity potential whose form is

$$
\begin{equation*}
\phi=\varepsilon \phi_{1}(x, t)+\varepsilon^{2} \phi_{2}(x, t)+\cdots \tag{1}
\end{equation*}
$$

The first-order term must satisfy the linear wave equation in Eq.(13.3.49), and the gradient of this term at $x=0$ must match the wall vibration. Thus, the first-order problem is

$$
\begin{gathered}
c^{2} \frac{\partial^{2} \phi_{1}}{\partial x^{2}}-\frac{\partial^{2} \phi_{1}}{\partial t^{2}}=0 \\
\left.\frac{\partial \phi_{1}}{\partial x}\right|_{x=0}=c \sin (\sigma t) \sin (\omega t)
\end{gathered}
$$

The velocity of the wall is a beat. We describe it in terms of complex exponentials. Doing so expedites solving the first-order equations, and it will also allow us to avoid recalling trigonometric identities for products when we go on to the second order. Thus, the boundary condition is converted to

$$
\begin{align*}
\left.\frac{\partial \phi_{1}}{\partial x}\right|_{x=0} & =-\frac{1}{4} \varepsilon c\left(e^{i \sigma t}-e^{-i \sigma t}\right)\left(e^{i \omega t}-e^{-i \omega t}\right)  \tag{2}\\
& =\frac{1}{4} \varepsilon c\left(e^{i(\omega-\sigma) t}-e^{i(\omega+\sigma) t}\right)+\text { c.c. }
\end{align*}
$$

This is a familiar problem. Each harmonic in the excitation generates a forward propagating wave at that frequency. Because there are two uncorrelated frequencies, it is preferable to use the retarded time to represent each term,

$$
\tau=t-\frac{x}{c}
$$

Thus, a suitable trial solution is

$$
\begin{equation*}
\phi_{1}=a_{1} e^{i \Omega_{1} \tau}+a_{2} e^{i \Omega_{2} \tau}+\text { c.c. } \tag{3}
\end{equation*}
$$

where

$$
\Omega_{1}=\omega-\sigma, \quad \Omega_{2}=\omega+\sigma
$$

Because $\partial \tau / \partial x=-1 / c$, substitution of this form into Eq. (2) yields

$$
\begin{equation*}
a_{1}=\frac{i c^{2}}{4 \Omega_{1}}, \quad a_{2}=-\frac{i c^{2}}{4 \Omega_{2}} \tag{4}
\end{equation*}
$$

We have seen that the moving boundary condition leads to second-order terms in $\phi$ and $p$ that are uniformly valid. Therefore, we may ignore that effect. Because the gradient of the first-order term matches the specified wall velocity, it must be that $\partial \phi_{2} / \partial x=0$ at $x=0$. The resulting second-order problem is

$$
\begin{align*}
& c^{2} \frac{\partial^{2} \phi_{2}}{\partial x^{2}}-\frac{\partial^{2} \phi_{2}}{\partial t^{2}}= 2(\beta-1) \frac{\partial^{2} \phi_{1}}{\partial x^{2}} \frac{\partial \phi_{1}}{\partial t}+2 \frac{\partial \phi_{1}}{\partial x} \frac{\partial^{2} \phi_{1}}{\partial x \partial t} \\
&\left.\frac{\partial \phi_{2}}{\partial x}\right|_{x=0}=0 \tag{5}
\end{align*}
$$

The differential equation is simplified by the properties that

$$
\frac{\partial \phi_{1}}{\partial t}=\frac{d \phi_{1}}{d \tau}, \quad \frac{\partial \phi_{1}}{\partial x}=-\frac{1}{c} \frac{d \phi_{1}}{d \tau}
$$

Consequently, the differential equation in Eq. (5) may be written as

$$
c^{2} \frac{\partial^{2} \phi_{2}}{\partial x^{2}}-\frac{\partial^{2} \phi_{2}}{\partial t^{2}}=2 \frac{\beta}{c^{2}} \frac{d \phi_{1}}{d \tau} \frac{d^{2} \phi_{1}}{d \tau^{2}} \equiv \frac{\beta}{c^{2}} \frac{d}{d \tau}\left[\left(\frac{d \phi_{1}}{d \tau}\right)^{2}\right]
$$

The term on the right side is formed by substituting for $\phi_{1}$ from Eq. (3). The use of c.c. in that expression is a convenience; care must be taken to account for the unlisted terms.

$$
\begin{align*}
& c^{2} \frac{\partial^{2} \phi_{2}}{\partial x^{2}}-\frac{\partial^{2} \phi_{2}}{\partial t^{2}}=\frac{\beta}{c^{2}} \frac{d}{d \tau}\left\{\left[i \Omega_{1} a_{1} e^{i \Omega_{1} \tau}+i \Omega_{2} a_{2} e^{i \Omega_{2} \tau}+\text { c.c. }\right]^{2}\right\} \\
&= \frac{\beta}{c^{2}} \frac{d}{d \tau}\left[-\Omega_{1}^{2} a_{1}^{2} e^{2 i \Omega_{1} \tau}-\Omega_{2}^{2} a_{2}^{2} e^{2 i \Omega_{2} \tau}-2 \Omega_{1} \Omega_{2} a_{1} a_{2} e^{2 i \omega \tau}\right. \\
&\left.+2 \Omega_{1} \Omega_{2} a_{1}^{*} a_{2} e^{2 i \sigma \tau}+\Omega_{1}^{2} a_{1} a_{1}^{*}+\Omega_{2}^{2} a_{2} a_{2}^{*}+\text { c.c. }\right]  \tag{6}\\
&=- \frac{2 i \beta}{c^{2}}\left[\Omega_{1}^{3} a_{1}^{2} e^{2 i \Omega_{1} \tau}+\Omega_{2}^{3} a_{2}^{2} e^{2 i \Omega_{2} \tau}\right. \\
&+\left.2 \omega \Omega_{1} \Omega_{2} a_{1} a_{2} e^{2 i \omega \tau}-2 \sigma \Omega_{1} \Omega_{2} a_{1}^{*} a_{2} e^{2 i \sigma \tau}\right]+ \text { c.c. }
\end{align*}
$$

The task now is to determine the portion of the particular solution for $\phi_{2}$ that grows as $x$ increases. Each term on the right side is solely a function of $\tau$, so each identically satisfies the homogeneous terms to the left of the equality side. Accordingly, we form a trial solution by multiplying each term by $x$. (The alternative of multiplying each term by $t$ must be rejected because the excitation is a periodic function of time, so the response must share that property.) Consequently, the trial form for $\left(\phi_{2}\right)_{\mathrm{p}}$ is

$$
\begin{equation*}
\left(\phi_{2}\right)_{\mathrm{p}}=x\left[b_{1} e^{2 i \Omega_{1} \tau}+b_{2} e^{2 i \Omega_{2} \tau}+b_{3} e^{2 i \omega \tau}+b_{4} e^{2 i \sigma t}\right]+\text { c.c. } \tag{7}
\end{equation*}
$$

Substitution of this ansatz into the left side of Eq. (6) generates terms that match the right side if

$$
\begin{gathered}
b_{1}=\frac{\beta}{2 c^{3}} \Omega_{1}^{2} a_{1}^{2}, \quad b_{2}=\frac{\beta}{2 c^{3}} \Omega_{2}^{3} a_{2}^{2} \\
b_{3}=\frac{\beta}{c^{3}} \Omega_{1} \Omega_{2} a_{1} a_{2}, \quad b_{4}=-\frac{\beta}{c^{3}} \Omega_{1} \Omega_{2} a_{1}^{*} a_{2}
\end{gathered}
$$

The complementary solution of Eq. (6) is the d'Alembert solution, which does not grow. Hence, we have completed the analysis. The velocity potential is constructed by substituting Eqs. (3) and (7) into Eq. (1), which yields

$$
\begin{align*}
\phi= & \varepsilon\left(a_{1} e^{i \Omega_{1} \tau}+a_{2} e^{i \Omega_{2} \tau}\right)+\varepsilon^{2}\left[b_{1}\left(x+\frac{c}{2 i \Omega_{1}}\right) e^{2 i \Omega_{1} \tau}+b_{2}\left(x+\frac{c}{2 i \Omega_{2}}\right) e^{2 i \Omega_{2} \tau}\right. \\
& \left.+b_{3}\left(x+\frac{c}{2 i \omega}\right) e^{2 i \omega \tau}+b_{4}\left(x+\frac{c}{2 i \sigma}\right) e^{2 i \sigma t}\right]+\operatorname{c.c}+O\left(\varepsilon^{2}\right) \tag{8}
\end{align*}
$$

This expression shows that the beating excitation at the boundary, which consists of two harmonic signals at close frequencies $\Omega_{1}=\omega-\sigma$ and $\Omega_{2}=\omega+\sigma$, generates double-frequency harmonics at $2(\omega+\sigma)$ and $2(\omega-\sigma)$, as well as a sum frequency harmonic at $2 \omega$, and a difference frequency harmonic at $2 \sigma$. All frequencies except $2 \sigma$ are close, so it seems that the waveform should display a beating pattern. These frequencies appear in both the particular and complementary solutions for $\phi_{2}$. The second-order terms grow with increasing $x$, so the representation is valid only for locations at which $x$ is small relative to some reference length, like the shock formation distance.

The expression we have derived does not provide much insight regarding the appearance of a waveform at substantial distances from the source. Although it was not requested, Fig. 1 shows the waveform obtained from the Riemann solution at $x_{\text {shock }}=c /(\beta \varepsilon \omega)$. The frequencies are $\omega /(2 \pi)=1000 \mathrm{~Hz}$ and $\sigma /(2 \pi)=40 \mathrm{~Hz}$. The envelope function that has been plotted is $\pm \varepsilon c \sin (\sigma t)$. This is the true envelope for linear theory, but a very close examination would show that the maxima and minima of each oscillation of $p$ sometimes do not reach the envelope, and sometimes it exceeds the envelope. These discrepancies are minor. On the other hand, the oscillation within the envelope shows the advancement and retardation that causes the waveform to lean over.


Figure 1.

Before we see how a nonuniformly valid perturbation series can be corrected, let us consider applying the Fourier series analysis in Sect. 13.3.2 to this problem. The condition satisfied by a period is that it contain an integer number of cycles at each frequency. In the case of the present excitation, it must be such that an interval $T$ contains $n$ oscillations at $\omega-\sigma$, and $m$ oscillations at $\omega+\sigma$. In other words, $T=n 2 \pi /(\omega-\sigma)=m 2 \pi /(\omega+\sigma)$. The smallest integers meeting this condition for the frequencies in Fig. 1 are $m=13$ and $n=12$, which corresponds to a period of $T=2 \pi / 80 \mathrm{~s}$. In other words, in a Fourier series analysis the frequencies contained in the excitation are the 12th and 13th harmonics of the fundamental frequency. The highest harmonic for the frequency series must be many times the frequency of the excitation if the propagation distance is comparable to $x_{\text {shock }}$. Thus, a very long series would be required if that formulation were to be applied to analyze the pressure generated by a beating vibration. The formulation that follows provides an alternative.

### 13.3.4 Method of Strained Coordinates

The occurrence of a nonuniformly valid description of a variable is not unique to analyses of nonlinearity. Consider a harmonic waveform whose frequency is slightly different from a nominal value $\omega_{0}$, for example, $y=\sin \left(\left(\omega_{0}+\varepsilon\right) t\right) \equiv$
$\sin \left(\omega_{0} t\right) \cos (\varepsilon t)+\cos \left(\omega_{0} t\right) \sin (\varepsilon t)$. This expression may be expanded in a Taylor series of increasing powers of $\varepsilon$. Truncation of that series at the first power gives $y \approx \sin \left(\omega_{0} t\right)+\varepsilon t \cos \left(\omega_{0} t\right)$. This is a nonuniformly valid perturbation series. Rather than considering the nonuniformity to stem from a small difference of the frequency from $\omega_{0}$, we could consider the frequency to be $\omega_{0}$ but the actual time variable to be slightly different from $t$. This is the essential idea underlying singular perturbation methods. Specifically, the nonuniformly valid perturbation series is regarded as the correct solution in the wrong place in space and/or time.

A widely employed method for correcting a nonuniformly valid perturbation series is Lighthill's method of strained coordinates. ${ }^{16}$ It distorts the space-time grid in a manner that is gradual when viewed over a small distance, but accumulates to a significant level at long range. If the nonuniformly valid term grows with increasing $x$, then the coordinate straining transformation replaces $x$ with a variable $\xi$ that differs from $x$ at a rate that matches the nonuniformly valid term(s), specifically,

$$
\begin{equation*}
x=\xi+\varepsilon x F(\xi, t) \tag{13.3.66}
\end{equation*}
$$

The concept is to introduce this transformation into an expression that is not uniformly valid and then select the function $F$ that regularizes the expression.

We will use the regular perturbation series for plane waves derived in the previous section as the framework for the development. The second-order velocity potential in Eq. (13.3.65) is not required to be uniformly valid because it is an abstract quantity that does not directly characterize the state of the fluid. Rather, the expressions for the pressure and particle velocity derived by differentiating $\phi$ must behave properly. An expression for $\phi$ is obtained by combining the first-order potential in Eq.(13.3.55) and the second-order potential in Eq.(13.3.65). It is permissible to ignore terms in $\phi$ that will lead to terms in $p$ and $v$ that are $O\left(\varepsilon^{2}\right)$ everywhere, because those terms will decrease in importance relative to the nonuniformly valid terms as the wave propagates. Therefore, we may begin with

$$
\begin{equation*}
\phi=-\varepsilon c X\left(t-\frac{x}{c}\right)-\varepsilon^{2} \frac{\beta}{2 c} x \dot{X}\left(t-\frac{x}{c}\right)^{2}+O\left(\varepsilon^{2}\right) \tag{13.3.67}
\end{equation*}
$$

A corollary of being allowed to drop terms that uniformly are $O\left(\varepsilon^{2}\right)$ or smaller is that we may use $p=-\rho_{0}(\partial \phi / \partial t)$. No approximation is contained in $v=\partial \phi / \partial x$. The state variables that result are

$$
\begin{gather*}
p=\varepsilon \rho_{0} c \dot{X}\left(t-\frac{x}{c}\right)+\varepsilon^{2} \rho_{0} \frac{\beta}{c} x \dot{X}\left(t-\frac{x}{c}\right) \ddot{X}\left(t-\frac{x}{c}\right)+O\left(\varepsilon^{2}\right) \\
v=\varepsilon \dot{X}\left(t-\frac{x}{c}\right)+\varepsilon^{2} \frac{\beta}{2 c}\left[\frac{2}{c} x \dot{X}\left(t-\frac{x}{c}\right) \ddot{X}\left(t-\frac{x}{c}\right)+\dot{X}\left(t-\frac{x}{c}\right)^{2}\right]+O\left(\varepsilon^{2}\right) \tag{13.3.68}
\end{gather*}
$$

[^75]The last term within the bracket for $v$ does not grow as $x$ increases. It is uniformly valid, so it may be placed in the group of terms that are uniformly $O\left(\varepsilon^{2}\right)$. Doing so leads to $p=\rho_{0} c v+O\left(\varepsilon^{2}\right)$, which is like the relation according to linear theory. This is not merely an interesting observation. Rather, it tells us that whatever we do to make $p$ uniformly valid will do the same for $v$.

We are now ready to introduce the coordinate straining transformation. We substitute the expression for $x$ in Eq. (13.3.66) into the preceding equation for $p$. The $O(\varepsilon)$ term is expanded in a Taylor series in powers of $\varepsilon$, according to

$$
\begin{align*}
\dot{X}\left(t-\frac{x}{c}\right) & =\dot{X}\left(t-\frac{\xi}{c}-\varepsilon \frac{x}{c} F(\xi, t)\right) \\
& =\dot{X}\left(t-\frac{\xi}{c}\right)-\varepsilon \frac{x}{c} F(\xi, t) \ddot{X}\left(t-\frac{\xi}{c}\right)+\cdots \tag{13.3.69}
\end{align*}
$$

Because $O\left(\varepsilon^{3}\right)$ terms are not being considered, the Taylor series may be truncated at the listed terms. For the same reason, we can simply replace $x$ with $\xi$ in the phase of the $O\left(\varepsilon^{2}\right)$ part of $\phi$. The result of these operations is

$$
\begin{align*}
p= & \varepsilon \rho_{0} c\left[\dot{X}\left(t-\frac{\xi}{c}\right)-\varepsilon \frac{x}{c} F(\xi, t) \ddot{X}\left(t-\frac{\xi}{c}\right)\right] \\
& +\varepsilon^{2} \rho_{0} \frac{\beta}{c} x \dot{X}\left(t-\frac{\xi}{c}\right) \ddot{X}\left(t-\frac{\xi}{c}\right)+O\left(\varepsilon^{2}\right) \tag{13.3.70}
\end{align*}
$$

The role of the $F$ function is to make the perturbation series regular. Thus, we select this function to cancel the growing term, that is

$$
\begin{equation*}
\rho_{0} c \varepsilon\left[-\varepsilon \frac{\xi}{c} F(\xi, t) \ddot{X}\left(t-\frac{\xi}{c}\right)\right]+\varepsilon^{2} \rho_{0}\left[\frac{\beta \xi}{c} \dot{X}\left(t-\frac{\xi}{c}\right) \ddot{X}\left(t-\frac{\xi}{c}\right)\right]=0 \tag{13.3.71}
\end{equation*}
$$

In other words

$$
\begin{equation*}
F(\xi, t)=\frac{\beta}{c} \dot{X}\left(t-\frac{\xi}{c}\right) \tag{13.3.72}
\end{equation*}
$$

The assembly of the results is

$$
\begin{gather*}
p=\rho_{0} c v+O\left(\varepsilon^{2}\right)=\rho_{0} c \varepsilon \dot{X}(\tau)+O\left(\varepsilon^{2}\right) \\
\tau=t-\frac{\xi}{c}, \quad x=\xi+\varepsilon \beta \frac{x}{c} \dot{X}(\tau) \tag{13.3.73}
\end{gather*}
$$

The second-order terms that remain in the expression for $p$ are those in Eq. (13.3.68) that do not grow with increasing $x$. They represent effects that always are small relative to the overall signal and therefore seldom are of interest. Some individuals refer to the process of removing the growing terms as a renormalization of the perturbation series. ("Normal" here refers to regularity, as opposed to singularity.)

Is the solution consistent with the Riemann solution? To show this to be true, we write the last of Eq. (13.3.73) as

$$
\begin{equation*}
\xi=x\left(1-\varepsilon \beta \frac{\dot{X}(\tau)}{c}\right) \tag{13.3.74}
\end{equation*}
$$

Correspondingly, the first and second of Eq. (13.3.73) are

$$
\begin{gather*}
p=\rho_{0} c v=\rho_{0} c \varepsilon \dot{X}(\tau)+O\left(\varepsilon^{2}\right) \\
\tau=t-\frac{x}{c}\left(1-\varepsilon \beta \frac{\dot{X}(\tau)}{c}\right)=t-\frac{x}{c}\left(1-\beta \frac{p}{\rho_{0} c^{2}}\right) \tag{13.3.75}
\end{gather*}
$$

For comparison recall Eq.(13.1.37), which is the small-signal approximation of the Riemann solution. The term containing $p$ in the denominator is small compared to one, so the denominator may be expanded in a binomial series. Doing so leads to

$$
\begin{equation*}
p=\rho_{0} c \dot{X}(\tau), \quad \tau=t-\frac{x}{c\left(1+\beta p /\left(\rho_{0} c^{2}\right)\right)}=t-\frac{x}{c}\left(1-\beta \frac{p}{\rho_{0} c^{2}}+O\left(\varepsilon^{2}\right)\right) \tag{13.3.76}
\end{equation*}
$$

For second order, this is the same as the expressions derived by the method of renormalization. It is useful to have this confirmation of the renormalization formulation, but for more general systems any confirmation must come from experiments.

Before we consider such generalizations, it is appropriate that we consider another singular perturbation technique that has been employed for nonlinear acoustics analyses. The method of multiple scales had been used in other areas for many years prior to Nayfeh and Kluwick's application ${ }^{17}$ of the method to analyze a finite amplitude plane wave. It recognizes that the spatial scale over which nonlinear effects become important is much greater than the scale of a wavelength. Each scale is defined to be an independent spatial variable $\xi_{n}$ that changes more slowly as $n$ increases. For the nonlinear wave equation, these variables are

$$
\begin{equation*}
\xi_{n}=\varepsilon^{n} x, \quad n=0,1, \ldots \tag{13.3.77}
\end{equation*}
$$

The velocity potential is considered to depend on each of these scaled variables, even though the value of each coordinate is set when $x$ and $\varepsilon$ are set. Thus, a derivative with respect to $x$ becomes

$$
\begin{equation*}
\frac{\partial \phi}{\partial x}=\frac{\partial \phi}{\partial \xi_{0}}+\varepsilon^{2} \frac{\partial \phi}{\partial \xi_{1}}+\varepsilon^{2} \frac{\partial \phi}{\partial \xi_{2}}+\cdots \tag{13.3.78}
\end{equation*}
$$

[^76]This transformation is introduced to the nonlinear wave equation and the boundary condition, after which the velocity potential is expanded in the perturbation series in Eq. (13.3.46). The first-order equations are

$$
\begin{gather*}
c^{2} \frac{\partial^{2} \phi_{1}}{\partial \xi_{0}^{2}}-\frac{\partial^{2} \phi_{1}}{\partial t^{2}}=0  \tag{13.3.79}\\
\left.\frac{\partial \phi_{1}}{\partial \xi_{0}}\right|_{x=0}=\dot{X}(t)
\end{gather*}
$$

and the second-order equations are

$$
\begin{gather*}
c^{2} \frac{\partial^{2} \phi_{2}}{\partial \xi_{0}^{2}}-\frac{\partial^{2} \phi_{2}}{\partial t^{2}}=-2 c^{2} \frac{\partial^{2} \phi_{1}}{\partial \xi_{0} \partial \xi_{1}}+2(\beta-1) \frac{\partial^{2} \phi_{1}}{\partial \xi_{0}^{2}} \frac{\partial \phi_{1}}{\partial t}+2 \frac{\partial \phi_{1}}{\partial \xi_{0}} \frac{\partial^{2} \phi_{1}}{\partial \xi_{0} \partial t} \\
\left.\frac{\partial \phi_{2}}{\partial \xi_{0}}\right|_{x=0}=-\left.\frac{\partial \phi_{2}}{\partial \xi_{0}}\right|_{x=0}-\left.X(t) \frac{\partial^{2} \phi_{1}}{\partial \xi_{0}^{2}}\right|_{x=0} \tag{13.3.80}
\end{gather*}
$$

Note that the second-order boundary condition accounts for the displaced location of the boundary, even though we know that effect does not lead to growing nonlinear effects.

Because we are not interested in $\phi$ terms that are $O\left(\varepsilon^{3}\right)$ or smaller, only the physical scale $\xi_{0}$ and the slow scale $\xi_{1}$ enter into the analysis. The first-order equations have the same form as those for a regular perturbation, but they allow $\phi_{1}$ to have arbitrary dependence on $\xi_{1}$. It is easier to satisfy boundary conditions after the general solution has been identified. Hence, the first-order solution is written as

$$
\begin{equation*}
\phi_{1}=f\left(t^{\prime}, \xi_{1}\right), \quad t^{\prime}=t-\frac{\xi_{0}}{c} \tag{13.3.81}
\end{equation*}
$$

This form is used to generate the second-order equations. Equation (13.3.57) may be used to describe the quadratic terms, so the second-order wave equation becomes

$$
\begin{equation*}
c^{2} \frac{\partial^{2} \phi_{2}}{\partial \xi_{0}^{2}}-\frac{\partial^{2} \phi_{2}}{\partial t^{2}}=2 c \frac{\partial}{\partial \xi_{1}}\left[\frac{\partial}{\partial t^{\prime}} f\left(t^{\prime}, \xi_{1}\right)\right]+\beta \frac{\partial}{\partial t^{\prime}}\left[\frac{\partial}{\partial t^{\prime}} f\left(t^{\prime}, \xi_{1}\right)\right]^{2} \tag{13.3.82}
\end{equation*}
$$

We know from the regular perturbation series analysis that the term whose coefficient is $\beta$ will lead to a particular solution for $\phi_{2}$ that is not uniformly valid. The role of the first term on the right side is to prevent this condition from occurring. That is, we set

$$
\begin{equation*}
\frac{\partial}{\partial \xi_{1}}\left[\frac{\partial}{\partial t^{\prime}} f\left(t^{\prime}, \xi_{1}\right)\right]=-\frac{\beta}{c} \frac{\partial}{\partial t^{\prime}}\left[\frac{\partial}{\partial t^{\prime}} f\left(t^{\prime}, \xi_{1}\right)\right]^{2} \tag{13.3.83}
\end{equation*}
$$

Fulfillment of this condition represents the primary difficulty in using the method of multiple scales. If considered in isolation, without prior knowledge of the nonlinear behavior of a plane wave, solution of this equation would not be a simple task. The result, after it is made to satisfy the first-order boundary condition, is

$$
\begin{gather*}
\phi_{1}=-c X(\tau)+\frac{\beta}{2 c} \xi_{1} \dot{X}(\tau)^{2}  \tag{13.3.84}\\
t^{\prime}=\tau-\frac{\beta}{c^{2}} \xi_{1} \dot{X}(\tau)
\end{gather*}
$$

From here, expressions for the pressure and particle velocity are obtained by differentiation of this function. We shall not pursue the details because the result will be equivalent to the Riemann solution. The primary motivation for this discussion of the method of multiple scales is to demonstrate that it will be substantially more difficult to implement than the method of strained coordinates. Such is especially the case when we consider situations where a signal depends on more than one spatial coordinate. Nevertheless, multiple scales method is a more general method than strained coordinates, so it is conceivable that only it would be suitable in some as yet unexplored situation.

### 13.4 Multidimensional Systems

Most linear acoustic systems of practical interest feature signals that are not truly planar waves. Higher-order modes in a rectangular waveguide are best characterized by Cartesian coordinates, whereas the signal within a cylindrical waveguide or radiated by a circular piston is best described in cylindrical coordinates. In the latter systems, the axial direction of the coordinate system would be the direction of propagation, whereas radiation from a cylinder features cylindrical coordinates with propagation in the transverse direction. Spherical coordinates are appropriate for radiation from a vibrating sphere. Nonlinear effects arise in each of these systems. Each has been successfully researched, some by analytical techniques, some by computational techniques, and some by both approaches. The scope of the investigations is too vast to survey here. An excellent starting point is the monograph edited by Hamilton and Blackstock, ${ }^{18}$ but even that work is merely an entry point to an enormous body of research. Our objective here is not to delve deeply into specific topics. Rather, we will consider a few systems that highlight methods and phenomena not seen in plane waves.

### 13.4.1 Finite Amplitude Spherical Wave

The choice of coordinate system to use for the analysis of any wave usually is dictated by the shape of the boundary. Here, we shall study how nonlinearity affects the propagation of a radially symmetric spherical wave. The signal is generated by a spherical body whose radius oscillates relative to a mean value $a$, such that $r=a+\varepsilon Y(t)$. The

[^77]magnitude of the radial velocity $\varepsilon \dot{Y}$ is much less than the speed of sound $c$, but it is sufficiently large to warrant consideration of nonlinear effects. In a sense, the signal is not a multidimensional wave because it depends on a single spatial coordinate, the radial distance $r$. Indeed, the basic line of investigation we shall pursue parallels the perturbation analysis of plane waves. However, dependence on the spherical coordinate $r$ rather than the Cartesian coordinate $x$ considerably complicates each stage of the analysis, and these complications are typical of those encountered in formulations using more general curvilinear coordinates. It is important to note that a much simpler analysis of spherical waves is available. ${ }^{19}$ The analysis here will show how the method of strained coordinates may be employed when simpler approaches are not available.

The radiated signal is radially symmetric, so $\phi$ is a function of radial distance $r$ and time $t$. The perturbation series for the velocity potential is

$$
\begin{equation*}
\phi=\varepsilon \phi_{1}(r, t)+\varepsilon^{2} \phi_{2}(r, t) \tag{13.4.1}
\end{equation*}
$$

The first-order equation derived from the quadratic nonlinear wave, Eq.(13.3.18), is

$$
\begin{equation*}
c^{2} \nabla^{2} \phi_{1}-\frac{\partial^{2} \phi_{1}}{\partial t^{2}}=0 \tag{13.4.2}
\end{equation*}
$$

We use this relation to replace $\nabla^{2} \phi_{1}$ in the second-order equation, which thereby is simplified to

$$
\begin{equation*}
c^{2} \nabla^{2} \phi_{2}-\frac{\partial^{2} \phi_{2}}{\partial t^{2}}=\frac{\partial}{\partial t}\left[\frac{(\beta-1)}{c^{2}}\left(\frac{\partial \phi_{1}}{\partial t}\right)^{2}+\nabla \phi_{1} \cdot \nabla \phi_{1}\right] \tag{13.4.3}
\end{equation*}
$$

[Equations (13.4.2) and (13.4.3) have been highlighted because they apply in any situation.] The operators in radially symmetric spherical coordinates are

$$
\begin{equation*}
\nabla^{2} \phi=\frac{1}{r} \frac{\partial^{2}}{\partial r^{2}}(r \phi), \quad \nabla \phi=\bar{e}_{r} \frac{\partial \phi}{\partial r} \tag{13.4.4}
\end{equation*}
$$

It is convenient to address satisfaction of the boundary condition at the surface, $r=a$, after the general solution has been obtained. The radiation condition requires that the signal be an outgoing wave, so we write the first-order solution as

$$
\begin{equation*}
\phi_{1}=\frac{a}{r} f\left(t^{\prime}\right) \tag{13.4.5}
\end{equation*}
$$

[^78]where the nature of the function $f$ is not specified at this juncture. The phase variable $t^{\prime}$ is the retarded time, which accounts for the time required for a signal to travel from the sphere's surface to the field point, that is,
\[

$$
\begin{equation*}
t^{\prime}=t-\frac{r-a}{c} \tag{13.4.6}
\end{equation*}
$$

\]

Because $\partial() / \partial t \equiv \partial() / \partial t^{\prime}$, substitution of the expression for $\phi_{1}$ into the secondorder wave equation leads to

$$
\begin{align*}
\frac{1}{r}\left[c^{2} \frac{\partial^{2}}{\partial r^{2}}\left(r \phi_{2}\right)-\frac{\partial^{2}}{\partial\left(t^{\prime}\right)^{2}}\left(r \phi_{2}\right)\right]= & \frac{\partial}{\partial t^{\prime}}\left[\frac{\beta a^{2}}{c^{2} r^{2}} \dot{f}\left(t^{\prime}\right)^{2}\right. \\
& \left.+\frac{2 a^{2}}{c r^{3}} f\left(t^{\prime}\right) \dot{f}\left(t^{\prime}\right)+\frac{a^{2}}{r^{4}} f\left(t^{\prime}\right)^{2}\right] \tag{13.4.7}
\end{align*}
$$

A consequence of the interchangeability of derivatives with respect to $t$ and $t^{\prime}$ is that an overdot will be used to denote differentiation with respect to the argument of $f$.

The preceding is a linear partial differential equation, so we may obtain its particular solution by superposing the contribution of each term on the right side. Each of those terms has the general form $r^{-n} Q\left(t^{\prime}\right)$. Let us examine the possibility that the particular solution has a similar form by considering $r \phi_{2}=z_{n}$, where

$$
\begin{equation*}
z_{n}=g(r) G\left(t^{\prime}\right) \tag{13.4.8}
\end{equation*}
$$

Then, the left side of Eq. (13.4.7) would be

$$
\begin{equation*}
\frac{1}{r}\left[c^{2} \frac{\partial^{2}}{\partial r^{2}} z_{n}-\frac{\partial^{2}}{\partial\left(t^{\prime}\right)^{2}} z_{n}\right]=-2 \frac{g^{\prime}(r)}{r} \dot{G}\left(t^{\prime}\right)+\frac{g^{\prime \prime}(r)}{r} G\left(t^{\prime}\right) \tag{13.4.9}
\end{equation*}
$$

If we set $g^{\prime}(r)=1 / r$, then $\dot{G}\left(t^{\prime}\right)$ can be selected to match the first term within the bracket of Eq. (13.4.7). However, doing so generates a term that is like the second term in the bracket because $g^{\prime \prime}(r)=-1 / r^{2}$. Suppose we add to $z_{n}$ another term for which $g^{\prime}(r)=1 / r^{2}$ in order to match the second term in the bracket plus the residual from the first term. Doing so generates a new residual term that resembles the third term in the bracket because it gives $g^{\prime \prime}(r)=-2 / r^{3}$. Addition of another term to $z_{n}$ for which $g^{\prime}(r)=1 / r^{3}$ would handle the $1 / r^{3}$ terms to the right, but also add another residual at $1 / r^{4}$. The process continues ad infinitum, with each added $z_{n}$ term selected to match the residual from the previous term. Hence, we conclude that we may construct a particular solution as an infinite series,

$$
\begin{gather*}
\left(r \phi_{2}\right)_{\mathrm{p}}=\sum_{n=1}^{\infty} g_{n}(r) G_{n}\left(t^{\prime}\right)  \tag{13.4.10}\\
g_{n}^{\prime}(r)=r^{-n}
\end{gather*}
$$

If we were to carry out the full analysis, we would be led to a recurrence relation wherein $\dot{G}_{n}\left(t^{\prime}\right)$ for $n>1$ is defined in terms of $G_{n-1}\left(t^{\prime}\right)$. Such an analysis would be required if it were necessary to identify all parts of $\phi_{2}$. Fortunately, this seldom is the case because of the nature of the $g_{n}(r)$ functions. Integration of the second of Eq. (13.4.10) leads to

$$
\begin{equation*}
g_{1}(r)=\ln (r), \quad g_{n}(r)=-\frac{1}{n-1} r^{1-n} \text { if } n>1 \tag{13.4.11}
\end{equation*}
$$

Hence, each term in Eq. (13.4.10) other than $n=1$ represents a contribution to $r \phi_{2}$ that decays at least as rapidly as $1 / r$. Because $r \phi_{1}$ does not decay with increasing $r$, all of those terms will be such that $\varepsilon^{2} \phi_{2}$ is much less than $\varepsilon \phi_{1}$ everywhere.

The same is not true for the $n=1$ term. Its contribution to $r \phi_{2}$ is proportional to $\ln (r)$. Although logarithmic growth is relatively slow, there is some large value of $r$ that would lead to $\varepsilon^{2} \phi_{2}$ having the same order of magnitude as $\varepsilon \phi_{1}$. This term represents the part that is responsible for the perturbation series not being uniformly valid.

Our interest is identification of terms that describe cumulative growth of nonlinear effects, so we modify Eq. (13.4.10) to be

$$
\begin{equation*}
r \phi_{2}=a \ln \left(\frac{r}{a}\right) G_{1}\left(t^{\prime}\right)+O\left(\frac{1}{r}\right) \tag{13.4.10}
\end{equation*}
$$

where the introduction of the factor $a$ will help avoid confusion regarding the dimensionality of quantities. Substitution of this expression converts the left side of Eq. (13.4.7) to

$$
\begin{equation*}
\frac{1}{r}\left[c^{2} \frac{\partial^{2}}{\partial r^{2}}\left(r \phi_{2}\right)_{\mathrm{p}}-\frac{\partial^{2}}{\partial\left(t^{\prime}\right)^{2}}\left(r \phi_{2}\right)_{\mathrm{p}}\right]=-\frac{2 a c}{r^{2}} \dot{G}_{1}\left(t^{\prime}\right) \tag{13.4.12}
\end{equation*}
$$

We equate this term to the first term on the right side of Eq.(13.4.7), which leads to

$$
\begin{equation*}
-\frac{2 a c}{r^{2}} G_{1}\left(t^{\prime}\right)=\frac{\beta a^{2}}{c^{2} r} \dot{f}\left(t^{\prime}\right)^{2} \tag{13.4.13}
\end{equation*}
$$

The terms on the right side of Eq.(13.4.7) that have not been matched would lead to $O(1 / r)$ and $O\left(1 / r^{2}\right)$ contributions to $r \phi_{2}$. Neither grows relative to $r \phi_{1}$, so these terms may be ignored. Thus, we have found that

$$
\begin{equation*}
\left(r \phi_{2}\right)_{\mathrm{p}}=-\frac{\beta a^{2}}{2 c^{3}} \ln \left(\frac{r}{a}\right) \dot{f}\left(t^{\prime}\right)^{2}+O\left(\frac{1}{r}\right) \tag{13.4.14}
\end{equation*}
$$

A corollary of this development is that there is no need to consider the complementary solution for $\phi_{2}$ because it is a solution of the linear wave equation. Such a term cannot lead to $\varepsilon^{2} \phi_{2}$ being comparable to $\varepsilon \phi_{1}$. In turn, because the role of the complementary solution is solely to satisfy the second-order boundary condition,
only the first-order condition need be satisfied. This is fortunate, because satisfying the second-order boundary condition, possibly accounting for the moving boundary that the surface is at $r=a+\varepsilon Y$, would require that we employ the full particular solution in Eq. (13.4.10).

The velocity potential obtained by substitution of the first- and second-order velocity potentials, Eqs. (13.4.5) and (13.4.14), into the perturbation series is

$$
\begin{equation*}
r \phi=\varepsilon a f\left(t^{\prime}\right)-\varepsilon^{2} \frac{\beta a^{2}}{2 c^{3}} \ln \left(\frac{r}{a}\right) \dot{f}\left(t^{\prime}\right)^{2}+O\left(\varepsilon^{2}\right) \tag{13.4.15}
\end{equation*}
$$

As a check, note that the dimensionality of $r \phi$ must be $\mathrm{L}^{3} / \mathrm{T}$. The $O(\varepsilon)$ term has those dimensions if $f$ is $\mathrm{L}^{2} / \mathrm{T}$, in which case the $O\left(\varepsilon^{2}\right)$ also is consistent. Identification of the coordinate straining transformation requires expressions for $p$ and $v_{r}$. Terms that are uniformly $O\left(\varepsilon^{2} / r\right)$ or less may be ignored, so we use $p=-\rho_{0} \partial \phi / \partial t^{\prime}$. In combination with $v_{r}=\partial \phi / \partial r$, we obtain

$$
\begin{gather*}
r p=-\rho_{0} \varepsilon a \dot{f}\left(t^{\prime}\right)+\rho_{0} \varepsilon^{2} \frac{\beta a^{2}}{c^{3}} \ln \left(\frac{r}{a}\right) \dot{f}\left(t^{\prime}\right) \ddot{f}\left(t^{\prime}\right)+O\left(\varepsilon^{2}\right)  \tag{13.4.16}\\
r v_{r}=-\varepsilon\left[\frac{a}{c} \dot{f}\left(t^{\prime}\right)(\tilde{t})+\frac{a}{r} f\left(t^{\prime}\right)\right]+\varepsilon^{2} \frac{\beta a^{2}}{c^{4}} \ln \left(\frac{r}{a}\right) \dot{f}\left(t^{\prime}\right) \ddot{f}\left(t^{\prime}\right)+O\left(\varepsilon^{2}\right)
\end{gather*}
$$

The loss of uniform validity for $r p$ and $r v_{r}$ is caused by a term that grows as $\ln (r / a)$ relative to the $O(\varepsilon)$ terms, so we shall strain the $r$ coordinate with a growth factor that is logarithmic. Thus, we try

$$
\begin{equation*}
r=\xi+\varepsilon a \ln \left(\frac{r}{a}\right) \mathcal{R} \tag{13.4.17}
\end{equation*}
$$

where $\mathcal{R}$ is an unknown function to be determined. The coordinate straining converts the retarded time to

$$
\begin{equation*}
t^{\prime} \equiv t-\frac{r-a}{c}=\tau-\varepsilon \frac{a}{c} \ln \left(\frac{r}{a}\right) \mathcal{R} \tag{13.4.18}
\end{equation*}
$$

The variable $\tau$ is a nonlinear retarded time. Its definition is

$$
\begin{equation*}
\tau=t-\frac{\xi-a}{c} \tag{13.4.19}
\end{equation*}
$$

The task now is to determine the function $\mathcal{R}$ that makes the expressions for $p$ and $v_{r}$ uniformly valid. Toward that end, Eq. (13.4.18) is used to replace $\tilde{t}$ as the argument of the $F$ and $\dot{F}$ functions. Taylor series expansions give

$$
\begin{align*}
& f\left(t^{\prime}\right)=f(\tau)-\varepsilon\left(\frac{a}{c}\right) \ln \left(\frac{r}{a}\right) \mathcal{R} \dot{f}(\tau)+O\left(\varepsilon^{2}\right) \\
& \dot{f}\left(t^{\prime}\right)=\dot{f}(\tau)-\varepsilon\left(\frac{a}{c}\right) \ln \left(\frac{r}{a}\right) \mathcal{R} \ddot{f}(\tau)+O\left(\varepsilon^{2}\right) \tag{13.4.20}
\end{align*}
$$

We substitute the transformation in Eq. (13.4.18) into Eq. (13.4.16). The result for $p$ is

$$
\begin{align*}
r p= & -\rho_{0} \varepsilon a\left[\dot{f}(\tau)-\varepsilon\left(\frac{a}{c}\right) \ln \left(\frac{r}{a}\right) \mathcal{R} \ddot{f}(\tau)\right] \\
& +\rho_{0} \varepsilon^{2} \frac{\beta a^{2}}{c^{3}} \ln \left(\frac{r}{a}\right) \dot{f}(\tau) \ddot{f}(\tau)+O\left(\frac{\varepsilon^{2} a}{r}\right) \tag{13.4.21}
\end{align*}
$$

Similar operations for $v_{r}$ give

$$
\begin{align*}
r v_{r} & =-\frac{\varepsilon}{c} a\left[\dot{f}(\tau)-\varepsilon\left(\frac{a}{c}\right) \ln \left(\frac{r}{a}\right) \mathcal{R} \ddot{f}(\tau)\right] \\
& -\varepsilon \frac{a}{r}\left[f(\tau)-\varepsilon\left(\frac{a}{c}\right) \ln \left(\frac{r}{a}\right) \mathcal{R} \dot{f}(\tau)\right]  \tag{13.4.22}\\
& +\varepsilon^{2} \frac{\beta a^{2}}{c^{4}} \ln \left(\frac{r}{a}\right) \dot{f}(\tau) \ddot{f}(\tau)+O\left(\varepsilon^{2}\right)
\end{align*}
$$

The role of $\mathcal{R}$ is to cancel all terms in $r p$ that have the factor $\varepsilon^{2} \ln (r / a)$. This condition occurs if

$$
\begin{equation*}
\mathcal{R}=-\frac{\beta}{c^{2}} \dot{f}(\tau) \tag{13.4.23}
\end{equation*}
$$

The pressure term that remains is

$$
\begin{equation*}
r p=-\rho_{0} \varepsilon a \dot{f}(\tau) \tag{13.4.24}
\end{equation*}
$$

The expression for $r v_{r}$ corresponding to the selection of $\mathcal{R}$ is

$$
\begin{equation*}
r v_{r}=-\frac{\varepsilon a}{c}\left[\dot{f}(\tau)+\frac{c}{r} f(\tau)\right]-\varepsilon^{2} \beta \frac{a^{2}}{c^{3} r} \ln \left(\frac{r}{a}\right) \dot{f}(\tau)^{2}+O\left(\varepsilon^{2}\right) \tag{13.4.25}
\end{equation*}
$$

The presence of $\ln (r / a)$ as a factor in the $O\left(\varepsilon^{2}\right)$ term might seem to mean that this expression is not uniformly valid. However, that factor is divided by $r$ and the limit as $r \rightarrow \infty$ of $\ln (r / a) / r$ is zero.

At this juncture, we have not addressed the determination of the $f$ function. It is somewhat less confusing to carry out that task if we work with a function that is dimensionless. Accordingly, we let

$$
\begin{equation*}
f(t)=-c^{2} F(t) \tag{13.4.26}
\end{equation*}
$$

This substitution converts the coordinate straining transformation to

$$
\begin{equation*}
r=\xi+\varepsilon \beta a \ln \left(\frac{r}{a}\right) \dot{F}(\tau) \tag{13.4.27}
\end{equation*}
$$

The corresponding nonlinear retarded time is

$$
\begin{equation*}
t^{\prime} \equiv t-\frac{r-a}{c}=\tau-\varepsilon \beta \frac{a}{c} \ln \left(\frac{r}{a}\right) \dot{F}(\tau) \tag{13.4.28}
\end{equation*}
$$

The corresponding uniformly valid expressions for the state variables are

$$
\begin{gather*}
p=\rho_{0} c^{2} \varepsilon \frac{a}{r} \dot{F}(\tau)+O\left(\frac{\varepsilon^{2} a}{r}\right) \\
v_{r}=\varepsilon c \frac{a}{r}\left[\dot{F}(\tau)+\frac{c}{r} F(\tau)\right]+\varepsilon^{2} c \beta \ln \left(\frac{r}{a}\right)\left(\frac{a^{2}}{r^{2}}\right) \dot{F}(\tau)^{2}+O\left(\frac{\varepsilon^{2} a}{r}\right) \tag{13.4.29}
\end{gather*}
$$

Aside from identification of $\tau$ dependence of $F$, the pressure has been determined. To see that this is so, suppose we specify $r$ and $t$, which sets $t-r / c$. Then, Eq. (13.4.28) is an algebraic equation for $\tau$. Substitution of the value of $\tau$ into the first of Eq. (13.4.29) gives the value of $p$ at that $r$ and $t$.

Determination of the $F$ function is done by satisfying the boundary condition on the sphere. At $r=a$, the coordinate straining gives $\tau=t$. To be consistent with the notation for plane waves, we let the specified velocity on the boundary be $\varepsilon c V(\tau)$. Velocity continuity at the surface requires that

$$
\begin{equation*}
\left.v_{r}\right|_{r=a}=\varepsilon c V(t) \tag{13.4.30}
\end{equation*}
$$

Equation (13.4.29) describes the particle velocity at $r=a$. We only need to substitute the $O(\varepsilon)$ terms, because the terms that are $O\left(\varepsilon^{2}\right)$ will merely lead to small corrections of $F(\tau)$. Thus, the substitution leads to a first-order differential equation for $F(\tau)$,

$$
\begin{equation*}
\dot{F}(\tau)+\frac{c}{a} F(\tau)=V(\tau) \tag{13.4.31}
\end{equation*}
$$

This is the same differential equation as that which must be solved in linear theory when the radial velocity is known at $r=a$. Indeed, the expressions for $p$ is the same as the description of the pressure for an outgoing wave according to linear theory, see Eq.(6.2.5), except that the retarded time has been replaced by $\tau$.

Before we proceed to an example, let us consider some qualitative aspects. Because nonlinear effects become progressively more significant at large $r$, it is reasonable to wonder whether they only occur in the farfield. A convenient reference distance marking where nonlinear effects are significant is that at which a shock forms. This is the closest location at which $\partial t / \partial \tau=0$ occurs at any instant. Differentiation of Eq. (13.4.28) gives

$$
\begin{equation*}
\frac{\partial t}{\partial \tau}=1-\varepsilon \beta \frac{a}{c} \ln \left(\frac{r_{\text {shock }}}{a}\right) \ddot{F}(\tau)=0 \tag{13.4.32}
\end{equation*}
$$

from which we find that

$$
\begin{equation*}
\ln \left(\frac{r_{\text {shock }}}{a}\right)=\frac{c / a}{\varepsilon \beta \max (\dddot{F})} \tag{13.4.33}
\end{equation*}
$$

The farfield of a spherical wave is characterized by $p$ being essentially $\rho_{0} c v_{r}$. To identify where this condition occurs let $T$ be a timescale representing the rate at which $F$ changes, that is, $O(|\dot{F}|)=O(|F|) / T$. We can assert from Eq. (13.4.29) that $p \approx \rho_{0} c v_{r}$ if $O(|F| / T) \gg(c / r) O(|F|)$, which is equivalent to $r \gg c T$. (If $a \gg c T$, then the farfield condition occurs everywhere.) Thus, the shock formation distance is inversely proportional to the maximum rate at which the pressure changes, whereas the farfield distance is independent of the pressure amplitude. This observation suggests that it might be possible to adjust the pressure to attain a shock at any range.

This hypothesis can be tested by considering the steady-state harmonic pressure corresponding to setting $\dot{f}(t)=\sin (\omega \tau)$. The steady-state nonlinear signal is described by

$$
\begin{equation*}
p=\varepsilon \rho_{0} c^{2}\left(\frac{a}{r}\right) \sin (\omega \tau) \tag{13.4.34}
\end{equation*}
$$

An appropriate reference time is $1 / \omega$, so the farfield occurs at $k r \gg 1$. The maximum value of $\ddot{f}$ is $\omega$, so the shock formation distance in this case is

$$
\begin{equation*}
r_{\text {shock }}=a e^{1 /(\varepsilon \beta k a)} \tag{13.4.35}
\end{equation*}
$$

If we arbitrarily define the farfield to occur when $k r>10$, we obtain

$$
\begin{equation*}
\frac{r_{\text {shock }}}{r_{\mathrm{ff}}}=\frac{k a}{10} e^{1 /\left(\varepsilon_{p} \beta k a\right)} \tag{13.4.36}
\end{equation*}
$$

The value of $\varepsilon$ is much less than one and $\beta$ is $O$ (1). Hence, it is evident that $r_{\text {shock }}$ will always be much greater than $r_{\mathrm{ff}}$. For example, a large value of $\varepsilon$ in air is 0.001 , in which case the above expression leads to $r_{\text {shock }}>170 r_{\text {ff }}$ for $k a<100$. A corollary of $r_{\text {shock }}$ being much greater than $r_{\mathrm{ff}}$ is that wherever a spherical wave is significantly affected by nonlinearity, it will be true that $p \approx \rho_{0} c v_{r}$.

Although the solution for a nonlinear spherical wave appears to be different from the Riemann solution, they actually are quite similar. To demonstrate the resemblance, we note that the first of Eq. (13.4.29) gives $\dot{F}(\tau)=(r / a) p /\left(\rho_{0} c^{2}\right)$. Substitution of this expression into Eq. (13.4.28) shows that the nonlinear retarded time is

$$
\begin{equation*}
\tau=t-\frac{r-a}{c}+\beta \frac{r}{c} \ln \left(\frac{r}{a}\right)\left(\frac{p}{\rho_{0} c^{2}}\right) \tag{13.4.37}
\end{equation*}
$$

For comparison, the characteristic variable for a weakly nonlinear plane wave was written in Eq. (13.3.76) as

$$
\begin{equation*}
\tau=t-\frac{x}{c}+\beta \frac{x}{c}\left(\frac{p}{\rho_{0} c^{2}}\right)+O\left(\varepsilon^{2}\right) \tag{13.4.38}
\end{equation*}
$$

The quantity $(r-a) / c$ in Eq. (13.4.37) is the time required for a linear signal to propagate from the surface to the field point. This time is analogous to $x / c$, which is the time required for a linear plane wave to travel from the boundary at which it is generated. One beneficial aspect of this similarity of the descriptions is that it is possible to adapt the earlier Fourier series analysis to handle situations where the boundary excitation is time-periodic. The adaptation is based on the fact that time occurs in the same manner in Eqs. (13.4.37) and (13.4.38). The implementation of this notion is the topic of Exercise 13.24.

The difference between linear and nonlinear retarded times for a spherical wave is proportional to $\ln (r / a)(r / c) p /\left(\rho_{0} c^{2}\right)$. In contrast, this difference for a planar wave is proportional to $(x / c) p /\left(\rho_{0} c^{2}\right)$. To assess the relative strength of the two effects, it is essential to bear in mind that spherical spreading leads to $r|p|$ being constant. Accordingly, we write the ratio of nonlinear effects as

$$
\begin{equation*}
\frac{\text { Spherical nonlinear term }}{\text { Planar nonlinear term }}=\frac{\ln (r / a)(r / c) p_{\text {spher }}}{(x / c) p_{\text {planar }}} \equiv \frac{a \ln (r / a)\left(r p_{\text {spher }}\right) / a}{x p_{\text {planar }}} \tag{13.4.39}
\end{equation*}
$$

If $p_{\text {spher }}$ at $r=a$ has the same order of magnitude as $p_{\text {planar }}$, then $\left(r p_{\text {spher }}\right) / a$ will have that magnitude at all $r$. This reduces the ratio of terms to $\ln (r / a) /(x / a)$. If $r$ and $x$ both equal some value $\ell \gg a$, then this ratio is much less than one. This observation leads to the conclusion that the nonlinear effect for a spherical wave is much weaker than it is for a plane wave.

Either of the computational algorithms in Sect. 13.1 .5 for evaluating profiles and waveforms of a planar signal may be adapted with small modification to evaluate spherical waves. For example, if we select a value of $\tau$, we may find $p$ from the first of Eq. (13.4.29). Then, Eq. (13.4.37) may be solved algebraically for $t$ if $r$ is specified. Solution for $r$ at a specified $t$ is more difficult because Eq. (13.4.37) is a transcendental function of $r$. For the same reason, determination of $p$ at specified $r$ and $t$ requires more effort. Like the case of planar waves, a nonlinear equation solver would be needed.

EXAMPLE 13.7 It is desired to compare the effect of nonlinearity in planar and spherical waves. Toward that end, consider a boundary motion in which the velocity normal to the surface is $\varepsilon c[\sin (\omega t)+(1 / 3) \sin (3 \omega t)]$. For the spherical wave, this is $v_{r}$ at $r=a$, whereas it is $v_{x}$ at $x=0$ for a plane wave propagating in the $x$ direction. Parameters are $\varepsilon=0.002, f=12 \mathrm{kHz}$, and $a=400 \mathrm{~mm}$. The fluid is water, for which $c=1480 \mathrm{~m} / \mathrm{s}$ and $\beta=3.5$. Determine the steady-state planar waveform at $x=x_{\text {shock }}$. Compare that result to the waveform of a spherical wave at $r=x_{\text {shock }}$ and at $r=r_{\text {shock }}$.

## Significance

Evaluation of the equations describing a nonlinear spherical wave is the primary emphasis. The results will vividly demonstrate the much slower advent of nonlinear effects that results from spherical spreading, as well as some surprising similarities to nonlinear distortion of plane waves.

## Solution

The function $F$ is obtained by solving the linear differential equation that is the boundary condition, Eq. (13.4.31). The specified boundary velocity is

$$
v_{r}=\varepsilon c V=\varepsilon c\left[\sin (\omega t)+\frac{1}{3} \sin (3 \omega t)\right]=\frac{\varepsilon c}{2 i}\left(e^{i \omega t}+\frac{1}{3} e^{3 i \omega t}\right)+\text { c.c. }
$$

The corresponding differential equation is

$$
\begin{equation*}
\dot{F}(\tau)+\frac{c}{a} F(\tau)=\frac{1}{2 i}\left(e^{i \omega \tau}+\frac{1}{3} e^{3 i \omega \tau}\right)+\text { c.c. } \tag{1}
\end{equation*}
$$

The form of the steady-state solution of Eq. (1) is

$$
\begin{equation*}
F(\tau)=\frac{1}{2 i}\left(B_{1} e^{i \omega \tau}+B_{3} e^{3 i \omega \tau}\right)+\text { c.c. } \tag{2}
\end{equation*}
$$

Substitution of this ansatz into Eq. (1) leads to expressions for the coefficients, which are found to be

$$
B_{1}=\frac{a / c}{1+i k a}, \quad B_{3}=\frac{a / c}{3+9 i k a}
$$

where $k=\omega / c$, as always. It is useful to work with a nondimensional version of $\dot{F}$, which is defined according to

$$
\begin{equation*}
F(\tau)=\frac{a}{c} \operatorname{Re}\left[\frac{1 / i}{1+i k a} e^{i \omega \tau}+\frac{1 / i}{3+9 i k a} e^{3 i \omega \tau}\right] \tag{3}
\end{equation*}
$$

Determination of $F$ allows us to finalize the expressions to evaluate. Equation (13.4.28) becomes

$$
\begin{equation*}
t-\frac{r-a}{c}=\tau-\varepsilon \beta \frac{a}{c} \ln \left(\frac{r}{a}\right) \operatorname{Re}\left[\frac{k a}{1+i k a} e^{i \omega \tau}+\frac{k a}{1+3 i k a} e^{3 i \omega \tau}\right] \tag{4}
\end{equation*}
$$

The pressure resulting from substitution of Eq. (2) into the first of Eq. (13.4.29) is

$$
\begin{equation*}
p=\rho_{0} c^{2} \varepsilon \frac{a}{r} \dot{\mathcal{F}}(\tau)=\rho_{0} c^{2} \varepsilon \frac{a}{r} \operatorname{Re}\left[\frac{k a}{1+i k a} e^{i \omega \tau}+\frac{k a}{1+3 i k a} e^{3 i \omega \tau}\right] \tag{5}
\end{equation*}
$$

The radial distance at which an arbitrary wave forms a shock is given by Eq. (13.4.33). In terms of the current functional representation, that expression is

$$
\begin{equation*}
\ln \left(\frac{r_{\text {shock }}}{a}\right)=\frac{c^{3}}{\varepsilon \beta a \max (-\ddot{F}(\tau))} \equiv \frac{1}{\varepsilon \beta k a \max (D(\tau))} \tag{6}
\end{equation*}
$$

where

$$
D(\tau)=\operatorname{Re}\left[i\left(\frac{i}{1+i k a} e^{i \omega \tau}+\frac{3 i}{1+3 i k a} e^{3 i \omega \tau}\right)\right]
$$

We could determine the maximum value of $D(\tau)$ by seeking the value of $\omega \tau$ at which $\dot{D}=0$. Such an evaluation is complicated by the fact that there is more than one local maximum. A simpler approach is to evaluate $D(\tau)$ for closely spaced values of $\omega \tau$ ranging over a period. The specified parameters give $k a=20.38$. A data scan gives max $(D(\tau))=1.9978$ at $\omega \tau /(2 \pi)=0.9984$. The shock formation distance obtained from Eq. (6) is $r_{\text {shock }}=13.37 \mathrm{~m}$.

We are now ready to evaluate waveforms. There is no requirement to compute the pressure values at a uniform time increment. Consequently, we may follow the simpler procedure in which we select $\omega \tau_{n}$ at equal increments in the range of a period $2 \pi$. The corresponding value of $p_{n}$ is obtained from Eq. (5). The instant $t_{n}$ at which $p_{n}$ occurs at a specified $r$ is found from Eq. (4), which gives

$$
\begin{equation*}
\omega t_{n}=k r-k a+\omega \tau-\varepsilon \beta k a \ln \left(\frac{r}{a}\right) \dot{F}(\tau) \tag{7}
\end{equation*}
$$

The plane wave also must be evaluated. The velocity at $x=0$ in the Riemann solution was written as $\varepsilon c V(t)$. The function $V(t)$ in the present situation is

$$
V(\tau)=\left.\frac{v_{x}}{\varepsilon c}\right|_{x=0}=\frac{1}{2 i}\left(e^{i \omega t}+\frac{1}{3} e^{3 i \omega t}\right)+\text { c.c. }
$$

The Riemann solution for the corresponding plane wave pressure is

$$
p=\varepsilon \rho_{0} c^{2} \operatorname{Re}\left(\frac{1}{i} e^{i \omega \tau}+\frac{1}{3 i} e^{3 i \omega \tau}\right), \omega t=\omega \tau+\frac{k x}{1+\varepsilon \beta V(\tau)}
$$

The shock formation distance for a plane wave is given by Eq. (13.2.26). For the present $V(\tau)$, the maximum value of $d V / d \tau$ is $2 \omega$ at $\omega \tau=0$. This leads to shock formation of the plane wave at

$$
x_{\text {shock }}=\frac{c}{2 \beta \varepsilon \omega}=1.402 \mathrm{~m}
$$

All computations may be carried out in a vectorial manner. Figure 1 compares the boundary velocity $V(\tau)$ to the pressure function $\dot{F}(\tau)$. Both are very close, and suggestive of a square wave. Both functions differ little because $k a$ is quite large. This means that the spherical wave exhibits its farfield behavior everywhere, so that $p \approx \rho_{0} c v_{r}$ at the surface of the sphere.


Figure 1.

The spherical waveforms at $r=x_{\text {shock }}$ and $r=r_{\text {shock }}$ in Fig. 2 have been delayed by the respective value of $(r-a) / c$, in order to place them in a common time base. It is difficult to compare them because of their vastly different magnitude. But it appears that the spherical wave is essentially undistorted at $r=x_{\text {shock }}$. This is not surprising because $r_{\text {shock }} \gg x_{\text {shock }}$.


Figure 2.

Multiplying each pressure waveform in Fig. 2 by their respective value of $r / a$ makes it possible to compare them and the plane wave at $x_{\text {shock }}$. This is the data in Fig. 3. Other than the spherical spreading factor, there is no perceptible difference between the spherical and plane waves at their respective shock formation distance. This behavior might seem to be surprising. It is attributable to the fact that $\ln (r / a)$ for a spherical wave and $x / x_{\text {shock }}$ for a plane wave are merely factors that scale the difference between the nonlinear and linear retarded times.


Figure 3.

To close this discussion, it is instructive to consider what would happen if $k a$ were much smaller. Reducing the frequency by a factor of ten with all other parameters constant leads to $r_{\text {shock }}=21.7\left(10^{15}\right) \mathrm{m}$. On the other hand, lowering the frequency by a factor of ten and raising the acoustic Mach number by a factor of ten change the shock formation distance by a relatively small amount, to $r_{\text {shock }}=18.86 \mathrm{~m}$. These trends show that nonlinear effects for a spherical wave are relatively small unless either the frequency, as measured by $k a$, or the acoustic Mach number is very large.

### 13.4.2 Waves in Cartesian Coordinates

There is no variation of pressure along the wavefronts of plane waves and radially symmetric spherical waves. Variation of pressure transverse to the direction of propagation gives rise to a phenomenon we have not yet encountered, specifically deviation of rays due to the signal associated with those rays. The system we shall investigate here exhibits this phenomenon, but it is somewhat academic. It concerns radiation from an infinite planar boundary, which is unrealizable. Nevertheless, with some modifications the solution can be used to describe the nonlinear propagation of signals in a hard-walled waveguide. ${ }^{20}$

Figure 13.21 depicts the system we will investigate. A flat plate extends infinitely in both directions. It is the boundary of the fluid in the infinite half-space above it. Knife-edge supports running perpendicularly to the plane of the diagram are spaced equidistantly at distance $L$. The plate vibrates in the direction normal to its surface, with the displacement $w$ defined to be positive into the fluid, which is the sense of increasing $z$. The displacement must be zero at its supports. The longest possible half-wavelength is $L$, so the displacement is set to

$$
\begin{equation*}
w=\frac{\varepsilon c}{\omega} \sin (\omega t) \sin \left(\frac{\pi x}{L}\right) \tag{13.4.40}
\end{equation*}
$$

(Shorter wavelengths along the plate may be studied by replacing $L$ with $L / n, n=$ $1,2, \ldots$ throughout the analysis.) We shall perform a steady-state analysis, so $w$ is presumed to have existed for a long time prior to $t=0$.


Fig. 13.21 A periodically supported vibrating plate, whose rest position is the plane $z=0$, generates a two-dimensional acoustic field

[^79]The task before us is to find a solution of the nonlinear wave equation such that the particle velocity and the velocity of the plate have equal components normal to the plate's surface. Another condition is that the acoustic signal must satisfy the Sommerfeld radiation condition, which requires that the pressure and particle velocity be outgoing or evanescent waves at large $z$. A further requirement follows from the fact that Fig. 13.21 does not change if we shift the coordinate system by any multiple of $2 L$ in either sense of the $x$ direction. Consequently, the solution must have period $2 L$ in the $x$ direction.

The condition of continuity of particle velocity at the fluid/plate interface constitutes a moving boundary condition. Eventually, the analysis will lead to a simplified boundary condition, as it did for plane and spherical waves. However, such an approximation might not be warranted in other systems, or in situations where nonlinear effects near the boundary are of interest. Hence, let us consider the "exact" condition for the plate. Figure 13.22 zooms in on a small portion of the surface at an arbitrary instant.


Fig. 13.22 Formulation of the continuity condition of particle velocity for fluid at the interface with a vibrating plate. The condition applies at the current position of the plate in the direction $\bar{n}$ normal to the current tangent to the plate

Figure 13.22 represents the displaced plate as the surface defined by $z=w(x, t)$. The normal to this surface is $\bar{n}=-\sin \chi \bar{e}_{x}+\cos \chi \bar{e}_{z}$, where $\chi$ is the angle between the normal to the plate and the $z$-axis. The particle velocity normal to the surface of the plate at its current position must equal the normal velocity of the plate at that position. In terms of the velocity potential, this condition is stated as

$$
\begin{equation*}
\left.\bar{n} \cdot \nabla \phi\right|_{z=w}=\bar{n} \cdot \frac{\partial w}{\partial t} \bar{e}_{z} \tag{13.4.41}
\end{equation*}
$$

Figure 13.22 shows that $\chi$ also is the angle between the tangent to the surface in the $x z$ plane and the $x$-axis, so $\chi=\tan ^{-1}(\partial w / \partial x)$. The displacement is $O(\varepsilon)$, so $|\chi| \ll 1$. Therefore, $\cos \chi \approx 1$ and $\sin \chi \approx \tan \chi \approx \partial w / \partial x$, so $\bar{n}=-(\partial w / \partial x) \bar{e}_{x}+\bar{e}_{z}$. When the gradient is replaced by the first two terms of a Taylor series relative to $z=0$, the boundary condition becomes

$$
\begin{equation*}
\left(-\frac{\partial w}{\partial x} \bar{e}_{x}+\bar{e}_{z}\right) \cdot\left[\left.\left.\nabla \phi\right|_{z=0}+w\left(\left.\frac{\partial}{\partial z} \nabla \phi \right\rvert\,\right)_{z=0} \right\rvert\,\right]_{z=0}=\frac{\partial w}{\partial t} \tag{13.4.42}
\end{equation*}
$$

The result of expanding this expression and dropping terms that are $O\left(\varepsilon^{3}\right)$ is

$$
\begin{equation*}
\frac{\partial \phi}{\partial z}-\frac{\partial w}{\partial x} \frac{\partial \phi}{\partial x}+w \frac{\partial^{2} \phi}{\partial z^{2}}=\frac{\partial w}{\partial t} \text { at } z=0 \tag{13.4.43}
\end{equation*}
$$

The first term is the one that would be employed for a linear analysis, the second compensates for the fact that $v_{x}$ has a normal component because $\bar{n}$ is rotated from the $z$-axis, and the third term compensates for the plate not being on the $x$-axis.

There is no variation of any feature of the system perpendicular to the plane of Fig. 13.21, so the velocity potential is a function of $x, z$, and $t$. The perturbation series is

$$
\begin{equation*}
\phi=\varepsilon \phi_{1}(x, z, t)+\varepsilon^{2} \phi_{2}(x, z, t)+\cdots \tag{13.4.44}
\end{equation*}
$$

Equations (13.4.2) and (13.4.3) respectively govern the first and second-order potentials.

The boundary conditions for the first and second-order potentials are found by substituting Eqs. (13.4.40) and (13.4.44) into the moving boundary condition and then matching like powers of $\varepsilon$. It is useful to use a complex exponential representation to handle products of harmonic time functions. Thus, we write

$$
\begin{align*}
\left.\frac{\partial \phi_{1}}{\partial z}\right|_{z=0} & =\frac{\partial}{\partial t}\left(\frac{w}{\varepsilon}\right)=\frac{1}{2} c e^{i \omega t} \sin \left(\frac{\pi x}{L}\right)+\text { c.c. } \\
\left.\frac{\partial \phi_{2}}{\partial z}\right|_{z=0} & =\left.\frac{\partial}{\partial x}\left(\frac{w}{\varepsilon}\right) \frac{\partial \phi_{1}}{\partial x}\right|_{z=0}-\left.\left(\frac{w}{\varepsilon}\right) \frac{\partial^{2} \phi_{1}}{\partial z^{2}}\right|_{z=0} \tag{13.4.45}
\end{align*}
$$

The sinusoidal dependence on $x$ on the boundary is like a higher-order mode in a two-dimensional waveguide. This suggests that the general solution for $\phi_{1}$ has the form

$$
\begin{equation*}
\phi_{1}=\frac{1}{2} B e^{i(\omega t-\kappa z)} \sin \left(\frac{\pi x}{L}\right)+\text { c.c. } \tag{13.4.46}
\end{equation*}
$$

This ansatz satisfies the first-order (linear) wave equation if the wavenumber is

$$
\begin{equation*}
\kappa=\left(k^{2}-\frac{\pi^{2}}{L^{2}}\right)^{1 / 2} \tag{13.4.47}
\end{equation*}
$$

If $k L<\pi$, the pressure will decay exponentially with increasing $z$. Such a situation is irrelevant to a study of nonlinearity at small $\varepsilon$ because the pressure will decay to insignificance before $\phi_{2}$ grows to a significant level. Thus, we shall only consider the case where the wavelength in the $z$ direction is less than $2 L$, that is, $k L>\pi$, which means that $\kappa$ is real. The coefficient $B$ is found by satisfying the first of Eq. (13.4.45). The resulting first-order potential is

$$
\begin{equation*}
\phi_{1}=\frac{1}{2}\left(\frac{i c}{\kappa} e^{i(\omega t-\kappa z)}+\text { c.c. }\right) \sin \left(\frac{\pi x}{L}\right)+\text { c.c. } \tag{13.4.48}
\end{equation*}
$$

To determine the second-order potential, we substitute $\phi_{1}$ into the second-order inhomogeneous wave equation, Eq. (13.4.3), which becomes

$$
\begin{align*}
c^{2} \nabla^{2} \phi_{2}-\frac{\partial^{2} \phi_{2}}{\partial t^{2}} & =\frac{1}{4} \frac{\partial}{\partial t}\left[\frac{(\beta-1)}{c^{2}}\left(-\frac{\omega c}{\kappa} e^{i(\omega t-\kappa z)}+\text { c.c. }\right)^{2} \sin \left(\frac{\pi x}{L}\right)^{2}\right. \\
& +\left(\frac{i c \pi}{\kappa L} e^{i(\omega t-\kappa z)}+\text { c.c. }\right)^{2} \cos \left(\frac{\pi x}{L}\right)^{2} \\
& \left.+\left(c e^{i(\omega t-\kappa z)}+\text { c.c. }\right)^{2} \sin \left(\frac{\pi x}{L}\right)^{2}\right] \tag{13.4.49}
\end{align*}
$$

The product of the complex exponential and its complex conjugate is a constant. There is no need to track such terms because each square is differentiated with respect to time. For the remaining terms we use the formulas for the cosine of a double angle to handle the $x$ dependence. Doing so leads to

$$
\begin{align*}
c^{2} \nabla^{2} \phi_{2}-\frac{\partial^{2} \phi_{2}}{\partial t^{2}}= & \frac{1}{8} \frac{\partial}{\partial t}\left\{e^{2 i(\omega t-\kappa z)}\left[\frac{(\beta-1) \omega^{2}}{\kappa^{2}}+c^{2}\right]\left[1-\cos \left(\frac{2 \pi x}{L}\right)\right]\right. \\
& \left.-\frac{\pi^{2} c^{2}}{\kappa^{2} L^{2}}\left[1+\cos \left(\frac{2 \pi x}{L}\right)\right]\right\}+ \text { c.c. } \tag{13.4.50}
\end{align*}
$$

The definition of $\kappa$ in Eq. (13.4.47) reduces the second-order equation to

$$
\begin{align*}
c^{2} \nabla^{2} \phi_{2}-\frac{\partial^{2} \phi_{2}}{\partial t^{2}} & =\frac{i \omega}{4 \kappa^{2}} e^{2 i(\omega t-\kappa z)}\left\{\left(\omega^{2} \beta-2 \frac{\pi^{2} c^{2}}{L^{2}}\right)\right. \\
& \left.-\omega^{2} \beta \cos \left(\frac{2 \pi x}{L}\right)\right\}+ \text { c.c. } \tag{13.4.51}
\end{align*}
$$

The second term is a solution of the linear wave equation. We can generate a particular solution by multiplying its form by $x, z$, and/or $t$. Other considerations dictate which of these alternatives should be chosen. Specifically, we know that the pressure must be periodic in $x$ and $t$ because shifting $t$ by $2 \pi / \omega$ or $x$ by $2 \pi / L$ does not alter the excitation. Factors of $t$ or $x$ multiplying sinusoidal terms would violate these periodicity requirements. Although the presence of a $z$ factor in $\phi_{2}$ might seem to violate the Sommerfeld radiation condition, the intent is to identify terms that lead to nonuniform validity. Thus, we shall try as a particular solution

$$
\begin{equation*}
\left(\phi_{2}\right)_{\text {part }}=\left[B_{0}+B_{1} z \cos \left(\frac{2 \pi x}{L}\right)\right] e^{2 i(\omega t-\kappa z)}+\text { c.c. } \tag{13.4.52}
\end{equation*}
$$

The coefficients are found by substituting this expression into Eq. (13.4.51) and then matching like terms. The result is

$$
\begin{equation*}
B_{0}=\frac{i c k}{16 \kappa^{2}} \frac{\beta k^{2}-2 \pi^{2} / L^{2}}{\left(k^{2}-\kappa^{2}\right)}, \quad B_{1}=\frac{c}{16} \beta\left(\frac{k^{3}}{\kappa^{3}}\right) \tag{13.4.53}
\end{equation*}
$$

The $B_{1}$ term leads to growth of $\phi_{2}$ relative to $\phi_{1}$ as $z$ increases, whereas the $B_{0}$ term remains bounded at all $z$. By definition, $\left(\phi_{2}\right)_{\text {comp }}$ is a solution of the (homogeneous) linear wave equation. As such, it too is bounded and therefore represents a contribution to the second-order potential that does not grow. For this reason, we need not determine it. (This is fortunate because substitution of $w$ and $\phi_{1}$ into the second of Eq. (13.4.45) would generate several terms that would need to be matched.) Consequently, the solution for the velocity potential reduces to

$$
\begin{equation*}
\phi=\varepsilon \frac{i c}{2 \kappa} e^{i(\omega t-\kappa z)} \sin \left(\frac{\pi x}{L}\right)+\varepsilon^{2} \frac{c \beta}{16}\left(\frac{k^{3}}{\kappa^{3}}\right) z e^{2 i(\omega t-\kappa z)} \cos \left(\frac{2 \pi x}{L}\right)+\text { c.c. }+O\left(\varepsilon^{2}\right) \tag{13.4.54}
\end{equation*}
$$

Although we could work with this representation, the presence of complex conjugate terms would complicate the process of identifying the coordinate straining transformation. The equivalent real form is

$$
\begin{align*}
\phi= & -\varepsilon \frac{c}{\kappa} \sin (\omega t-\kappa z) \sin \left(\frac{\pi x}{L}\right)  \tag{13.4.55}\\
& +\varepsilon^{2} \frac{c \beta}{8}\left(\frac{k^{3}}{\kappa^{3}}\right) z \cos (2 \omega t-2 \kappa z) \cos \left(\frac{2 \pi x}{L}\right)+O\left(\varepsilon^{2}\right)
\end{align*}
$$

As always, $O\left(\varepsilon^{2}\right)$ should be understood to denote terms that everywhere have that order of magnitude.

The nonuniform validity of this expression is not addressed directly, because the velocity potential is not a state variable for the state of the fluid. The fact that $\phi$ is not uniformly valid is passed on to the pressure and particle velocity, which are state variables. However, the growth factor will be different. This is a two-dimensional system, so there are two particle velocity components, given by

$$
\begin{align*}
v_{x}=\frac{\partial \phi}{\partial x}= & \varepsilon c\left[-\left(\frac{\pi}{\kappa L}\right) \sin (\omega t-\kappa z) \cos \left(\frac{\pi x}{L}\right)\right. \\
& \left.-\varepsilon \frac{\beta}{4}\left(\frac{\pi}{L}\right)\left(\frac{k^{3}}{\kappa^{3}}\right) z \cos (2 \omega t-2 \kappa z) \sin \left(\frac{2 \pi x}{L}\right)\right]+O\left(\varepsilon^{2}\right) \\
v_{z}=\frac{\partial \phi}{\partial z}= & \varepsilon c\left[\cos (\omega t-\kappa z) \sin \left(\frac{\pi x}{L}\right)\right. \\
& \left.+\varepsilon \frac{\beta}{4}\left(\frac{k^{3}}{\kappa^{2}}\right) z \sin (2 \omega t-2 \kappa z) \cos \left(\frac{2 \pi x}{L}\right)\right]+O\left(\varepsilon^{2}\right) \tag{13.4.56}
\end{align*}
$$

The pressure may be evaluated according to $p=-\rho_{0} \partial \phi / \partial t$ because the quadratic terms in Eq. (13.3.20), being formed from $\phi_{1}$, would merely add $O\left(\varepsilon^{2}\right)$ terms that do not grow relative to the $O(\varepsilon)$ terms. Because time only occurs in $\phi$ as $\omega t-\kappa z$, we have $(\partial \phi / \partial t) / \omega=-(\partial \phi / \partial x) / \kappa+O\left(\varepsilon^{2}\right)$, from which it follows that

$$
\begin{equation*}
p=\left(\rho_{0} \omega / \kappa\right) v_{z}+O\left(\varepsilon^{2}\right) \tag{13.4.57}
\end{equation*}
$$

The descriptions of $p, v_{x}$, and $v_{z}$ that have been derived are not uniformly valid. However, because $p$ is proportional to $v_{z}$, it is only necessary to find a coordinate straining transformation by considering $v_{x}$ and either $v_{z}$ or $p$. Both $x$ and $z$ should be transformed because two state variables require correction, and the second-order term for $v_{x}$ and $v_{z}$ is different. A different perspective is that the velocity components are comparable in magnitude, so we should anticipate that the spatial grid should be strained comparably in both directions. The growth factor is proportional to $z$, so let us try

$$
\begin{align*}
& x=\eta+\varepsilon z F(\eta, \xi, t) \\
& z=\xi+\varepsilon z G(\eta, \xi, t) \tag{13.4.58}
\end{align*}
$$

We substitute the coordinate transformation into the real forms of the state variables, which gives

$$
\begin{gather*}
\frac{v_{x}}{\varepsilon c}=-\left(\frac{\pi}{\kappa L}\right) \sin [\omega t-\kappa(\xi+\varepsilon z G(\eta, \xi, t))] \cos \left[\frac{\pi}{L}(\eta+\varepsilon z F(\eta, \xi, t))\right] \\
-\varepsilon \frac{\beta}{4}\left(\frac{\pi}{L}\right)\left(\frac{k^{3}}{\kappa^{3}}\right) z \cos (2 \omega t-2 \kappa \xi) \sin \left(\frac{2 \pi \eta}{L}\right)+O(\varepsilon) \\
\frac{v_{z}}{\varepsilon c}=\frac{\kappa}{k}\left(\frac{p}{\varepsilon \rho_{0} c^{2}}\right)=\cos [\omega t-\kappa(\xi+\varepsilon z G(\eta, \xi, t))] \sin \left[\frac{\pi}{L}(\eta+\varepsilon z F(\eta, \xi, t))\right]  \tag{13.4.59a}\\
+\varepsilon \frac{\beta}{4} \frac{k^{3}}{\kappa^{2}} z \sin (2 \omega t-2 \kappa \xi) \cos \left(\frac{2 \pi \eta}{L}\right)+O(\varepsilon) \tag{13.4.59b}
\end{gather*}
$$

The terms containing $F$ or $G$ are expanded in Taylor series that are truncated at $O(\varepsilon)$. In addition, the double angle formulas are applied to the $O\left(\varepsilon^{2} z\right)$ terms,

$$
\begin{aligned}
\frac{v_{x}}{\varepsilon c}= & -\left(\frac{\pi}{\kappa L}\right)\left[\sin (\omega t-\kappa \xi) \cos \left(\frac{\pi \eta}{L}\right)\right. \\
& -(\kappa \varepsilon z G) \cos (\omega t-\kappa \xi) \cos \left(\frac{\pi \eta}{L}\right) \\
& \left.-\left(\frac{\pi}{L} \varepsilon z F\right) \sin (\omega t-\kappa \xi) \sin \left(\frac{\pi \eta}{L}\right)\right] \\
& -\varepsilon \frac{\beta}{2}\left(\frac{\pi}{L}\right)\left(\frac{k^{3}}{\kappa^{3}}\right) z\left[\cos (\omega t-\kappa \xi)^{2}\right. \\
& \left.-\sin (\omega t-\kappa \xi)^{2}\right] \sin \left(\frac{\pi \eta}{L}\right) \cos \left(\frac{\pi \eta}{L}\right)+O\left(\varepsilon^{2}\right) \\
\frac{v_{z}}{\varepsilon c}=\frac{\kappa}{k}\left(\frac{p}{\varepsilon \rho_{0} c^{2}}\right)= & \cos (\omega t-\kappa \xi) \sin \left(\frac{\pi \eta}{L}\right)+(\kappa \varepsilon z G) \sin (\omega t-\kappa \xi) \sin \left(\frac{\pi \eta}{L}\right) \\
& +\left(\frac{\pi}{L} \varepsilon z F\right) \cos (\omega t-\kappa \xi) \cos \left(\frac{\pi \eta}{L}\right)
\end{aligned}
$$

$$
\begin{align*}
& \quad+\varepsilon \frac{\beta}{2} \frac{k^{3}}{\kappa^{2}} z \sin (\omega t-\kappa \xi) \cos (\omega t \\
& -\kappa \xi)\left[\cos \left(\frac{\pi \eta}{L}\right)^{2}-\sin \left(\frac{\pi \eta}{L}\right)^{2}\right]+O\left(\varepsilon^{2}\right) \tag{13.4.60}
\end{align*}
$$

The task now is to determine $F$ and $G$ functions that remove all terms in both $v_{x}$ and $v_{z}$ that have $z$ as a factor. A clue suggesting the appropriate choice comes from recognition that in both expressions, the sinusoidal factors of $G$ are shifted by $90^{\circ}$ relative to those for $F$. This suggests that $G$ also should be $90^{\circ}$ out-of-phase relative to $F$ in both its $\xi$ and $\eta$ dependencies. Each should match the nonuniformly valid term derived from $\phi$. Thus, we try

$$
\begin{align*}
F & =C_{F} \sin (\omega t-\kappa \xi) \cos \left(\frac{\pi \eta}{L}\right)  \tag{13.4.61}\\
G & =C_{G} \cos (\omega t-\kappa \xi) \sin \left(\frac{\pi \eta}{L}\right)
\end{align*}
$$

These forms are substituted into Eq.(13.4.60), and terms having like dependence on $\xi$ and $\eta$ are grouped. The values of $C_{F}$ and $C_{G}$ are those that result in annihilation of all terms that have $z$ as a factor. This leads to

$$
\begin{equation*}
C_{F}=-\frac{\beta k L}{2 \pi}\left(\frac{k}{\kappa}\right)^{2}, \quad C_{G}=\frac{\beta}{2}\left(\frac{k}{\kappa}\right)^{3} \tag{13.4.62}
\end{equation*}
$$

The corresponding coordinate straining transformation is

$$
\begin{gather*}
x=\eta-\frac{1}{2} \varepsilon \beta z \frac{k L}{\pi}\left(\frac{k}{\kappa}\right)^{2} \sin (\omega t-\kappa \xi) \cos \left(\frac{\pi \eta}{L}\right)  \tag{13.4.63}\\
z=\xi+\frac{1}{2} \varepsilon \beta z\left(\frac{k}{\kappa}\right)^{3} \cos (\omega t-\kappa \xi) \sin \left(\frac{\pi \eta}{L}\right)
\end{gather*}
$$

The terms that remain in the state variables are

$$
\begin{gather*}
\frac{v_{x}}{\varepsilon c}=-\frac{\pi}{\kappa L} \sin (\omega t-\kappa \xi) \cos \left(\frac{\pi \eta}{L}\right)+O\left(\varepsilon^{2}\right)  \tag{13.4.64}\\
\frac{v_{z}}{\varepsilon c}=\frac{\kappa}{k}\left(\frac{p}{\varepsilon \rho_{0} c^{2}}\right)=\cos (\omega t-\kappa \xi) \sin \left(\frac{\pi \eta}{L}\right)+O\left(\varepsilon^{2}\right)
\end{gather*}
$$

Evaluation of the pressure corresponding to specified $x, z$, and $t$ requires that Eq.(13.4.63) be solved for $\eta$ and $\xi$. Such a determination will require numerical methods because of the transcendental nature of the transformation. These matters will be addressed in the next example. Here, we shall focus on the meaning of the coordinate straining. The general concept of a coordinate straining transformation is that it shifts the location where the linear signal is observed. The linear solution
is recovered by setting $\varepsilon=0$ in Eq.(13.4.63), which leads to $\xi=z$ and $\eta=x$ in Eq. (13.4.64). Thus, one way to view the nonlinear effect is to view lines of constant $\xi$ and constant $\eta$ as a grid. Such a view is afforded by Fig. 13.23, which shows the transformation grid at two instants separated by a quarter-period. (The value of $\varepsilon$ used there is quite large because doing so assists visualizations.)


Fig. 13.23 Instantaneous visualizations of the coordinate straining transformation for the nonlinear signal radiated by a vibrating flat plate, $\varepsilon=0.005, k L=3.628$. Constant $\xi$ lines are solid and red; constant $\eta$ lines are dashed and blue

The phase variable for the signal is

$$
\begin{equation*}
\theta=\omega t-\kappa \xi \tag{13.4.65}
\end{equation*}
$$

The coordinate transformation gives $\xi=0$ at $z=0$, so all signals on the wavefront at specific values of $t$ and $\xi$ left the plate at $t_{0}=t-(\kappa / \omega) \xi$. Thus, lines of constant $\xi$ are wavefronts. The transformation also states that $\eta=x$ at $z=0$. The signal corresponding to specific $\xi$ and $\eta$ at time $t$ left the plate from location $x=\eta$ at time $t-(\kappa / \omega) \xi$. Thus, a line of constant $\eta$ is the locus of locations at instant $t$ of signals that departed from $x=\eta$ prior to time $t$. (These lines were mistakenly referred to as rays in prior publications describing this system. ${ }^{21}$ The rays are identified later in this discussion). Figure 13.23 shows that the constant $\xi$ and $\eta$ lines near the plate differ little from the linear grid, $z=\eta$ and $\xi=\eta$. The differences grow in an oscillatory manner with increasing $z$, but the lines $\xi= \pm L / 2$ are straight.

[^80]Each of these behaviors has a relatively simple physical explanation. Note that the combination of sinusoidal terms in each of Eq. (13.4.63) matches a velocity component in Eq. (13.4.64). When we use this fact to eliminate the sinusoidal terms in the coordinate straining, we see that

$$
\begin{gather*}
x=\eta+\frac{1}{2} \beta z\left(\frac{k L}{\pi}\right)^{2}\left(\frac{k}{\kappa}\right)\left(\frac{v_{x}}{c}\right)  \tag{13.4.66}\\
z=\xi+\frac{1}{2} \beta z\left(\frac{k}{\kappa}\right)^{3}\left(\frac{v_{z}}{c}\right)
\end{gather*}
$$

A signal corresponding to specific $\xi$ and $\eta$ would occur at $x=\xi$ and $x=\eta$ in the absence of nonlinearity. The nonlinear signal occurs at a location that is shifted in the $x$ direction by an amount that is proportional to its $v_{x}$ value. That location also is shifted in the $z$ direction by an amount that is proportional to its $v_{z}$ value. In other words, the manner in which a signal propagates is altered by the signal. The constant $\xi$ lines in Fig. 13.23 are oscillatory because $v_{z}$ depends sinusoidally on $\eta$. Similarly, the oscillatory nature of the constant $\eta$ lines is due to the sinusoidal dependence of $v_{x}$ on $\eta$. The lines $\eta= \pm L / 2$ are straight because $v_{x}$ is zero for signals that may be traced back to the plate at $= \pm L / 2$.

Advancement and retardation of a wavefront is the amplitude dispersion process encountered in plane and spherical waves. Indeed, if we set $\kappa=k$, the difference between $z$ and $\xi$ is half the value for a plane wave. This process is weaker here because $v_{z}$ fluctuates along the wavefront. Shifting of the signal in the direction transverse to the propagation due to the signal's particle velocity in that direction is the process of self-refraction. It does not occur in plane waves. We will see in the next example that the presence of both processes sometimes leads to waveform distortion that is unlike what we have observed thus far.

The zoomed view in Fig. 13.23 shows that at large $z$ the grid lines for different $\xi$ and $\eta$ have regions where they intersect or are out of synchronization. For example, a wavefront for a lesser value of $\xi$ might be at a larger value of $z$. This, of course, corresponds to multivaluedness of the signal. In other words, a shock is previously formed as some distance closer to the boundary. One way of identifying a shock is to extend the approach for plane waves in which the transformation is analyzed. Suppose we look at the signal at a point $x, z$ that corresponds to strained coordinates $\eta, \xi$. If we increment the latter by differential amounts $d \eta$ and $d \xi$, we will obtain the signal at $x+d x, z+d z$. Let us write Eq. (13.4.63) generically as $x=f_{x}(\eta, \xi, t, z), x=$ $f_{x}(\eta, \xi, t, z)$. Then the differential increments are related to

$$
\begin{align*}
& d x=d \eta \frac{\partial f_{x}}{\partial \eta}+d \xi \frac{\partial f_{x}}{\partial \xi} \\
& d z=d \eta \frac{\partial f_{z}}{\partial \eta}+d \xi \frac{\partial f_{z}}{\partial \xi} \tag{13.4.67}
\end{align*}
$$

The matrix representation of these relations is

$$
\left\{\begin{array}{l}
d x  \tag{13.4.68}\\
d z
\end{array}\right\}=\left[\begin{array}{ll}
\frac{\partial f_{x}}{\partial \eta} & \frac{\partial f_{x}}{\partial \xi} \\
\frac{\partial f_{z}}{\partial \eta} & \frac{\partial f_{z}}{\partial \xi}
\end{array}\right]\left\{\begin{array}{l}
d \eta \\
d \xi
\end{array}\right\}
$$

The coefficient matrix is the Jacobian of the transformation. If its determinant is zero, the transformation is not well-defined. This is the condition for the advent of a shock. Expansion of the determinant and the formula for the cosine of the sum of two angles show that either of the two conditions leads to the Jacobian being zero,

$$
\begin{align*}
& \frac{1}{2} \varepsilon \beta(k z)\left(\frac{k}{\kappa}\right)^{2} \cos \left(\omega t-\kappa \xi+\frac{\pi}{L} \eta\right)=1 \\
& \frac{1}{2} \varepsilon \beta(k z)\left(\frac{k}{\kappa}\right)^{2} \cos \left(\omega t-\kappa \xi-\frac{\pi}{L} \eta\right)=-1 \tag{13.4.69}
\end{align*}
$$

Either relation in combination with Eq. (13.4.63) constitutes three equations relating the values of $\eta, \xi, x$, at a specified $t$. Their simultaneous solution for fixed $t$ defines the locus of points at which a shock forms at that instant.

The views in Fig. 13.23 provide an instantaneous picture of the location of signals that were radiated from the plate in the prior interval. In some respects, a more meaningful question is: What is the path followed by the signal that radiated from the plate at a specific instant? By definition, this path is a ray. All properties of the signal on a ray other than its amplitude are constant. Consequently, a ray in the present system is the locus of points at which a constant value of the phase variable $\theta=\omega t-\kappa \xi$ and strained coordinate $\eta$ is observed as time elapses.

In linear theory, $\eta=x$ and $\xi=z$. Constant $\theta$ in that case is observed at $z=(\omega t-\theta) / \kappa$ on a line of constant $x$. Thus, the rays of a linearized description are perpendicular to the plate, and the signal propagates along the ray at phase speed $\omega / \kappa$. The situation is different when nonlinear effects are included. Identification of rays and wavefronts begins by noting that the strained coordinate corresponding to designated values of $\theta$ and $t$ is $\xi=(\omega t-\theta) / \kappa$. This relation is used to eliminate $\xi$ from Eq. (13.4.63). Then the equation is solved for $x$ and $z$ as functions of $\theta, t$, and $\eta$. The second equation gives

$$
\begin{align*}
z & =\frac{\omega t-\theta}{\kappa\left[1-\frac{1}{2} \varepsilon \beta\left(\frac{k^{3}}{\kappa^{3}}\right) \cos (\theta) \sin \left(\frac{\pi \eta}{L}\right)\right]}  \tag{13.4.70}\\
& =\left(\frac{\omega t-\theta}{\kappa}\right)\left[1+\frac{1}{2} \varepsilon \beta\left(\frac{k^{3}}{\kappa^{3}}\right) \cos (\theta) \sin \left(\frac{\pi \eta}{L}\right)\right]+O\left(\varepsilon^{2}\right)
\end{align*}
$$

This expression is used to eliminate $z$ in the first of Eq. (13.4.63), with the result that

$$
\begin{equation*}
x=\eta-\left(\frac{\omega t-\theta}{\kappa}\right) \varepsilon \beta \frac{k L}{2 \pi}\left(\frac{k}{\kappa}\right)^{2} \sin (\theta) \cos \left(\frac{\pi \eta}{L}\right)+O\left(\varepsilon^{2}\right) \tag{13.4.71}
\end{equation*}
$$

To obtain a ray, we fix $\theta$ and $\eta$, and consider $t$ to be a variable parameter. Equations (13.4.70) and (13.4.71) give the $x$ and $z$ coordinates of the point at which this phase is situated at a specific $t$. Both equations are linear in $t$, so it follows that a ray is a straight line. The slope of this line relative to the $z$-axis is

$$
\begin{equation*}
\left(\frac{d x}{d z}\right)_{\text {ray }}=\frac{\partial x / \partial t}{\partial z / \partial t}=-\varepsilon \beta \frac{k L}{2 \pi}\left(\frac{k^{2}}{\kappa^{2}}\right) \sin (\theta) \cos \left(\frac{\pi \eta}{L}\right)+O\left(\varepsilon^{2}\right) \tag{13.4.72}
\end{equation*}
$$

Self-refraction is evidenced here by the fact that the sinusoidal terms match $v_{x}$ in Eq. (13.4.64), which leads to

$$
\begin{equation*}
\left(\frac{d x}{d z}\right)_{\mathrm{ray}}=\frac{1}{2} \beta\left(\frac{k L}{\pi}\right)^{2}\left(\frac{k}{\kappa}\right)^{2}\left(\frac{v_{x}}{c}\right) \tag{13.4.73}
\end{equation*}
$$

The value of $\varepsilon$ is small, so the slope essentially is the angle. From this we conclude that a ray is rotated from perpendicularity to the $z=0$ plane by an amount that is proportional to the particle velocity $v_{x}$ in the signal associated with that ray.

To determine the phase speed at which a signal propagates along its ray as $t$ elapses, we observe that if $\theta$ and $\eta$ are constant, then Eq. (13.4.70) is an explicit expression for $z$. Differentiating it gives

$$
\begin{equation*}
\left.\frac{d z}{d t}\right|_{\text {constant } \theta \text { and } \eta}=\left(\frac{\omega}{\kappa}\right)\left[1+\frac{1}{2} \varepsilon \beta\left(\frac{k^{3}}{\kappa^{3}}\right) \cos (\theta) \sin \left(\frac{\pi \eta}{L}\right)\right] \tag{13.4.74}
\end{equation*}
$$

The actual phase velocity is parallel to the ray; the preceding is the $z$ component of that velocity. However the angle between the ray and the $z$-axis is $O(\varepsilon)$, so the $z$ component is essentially the magnitude of the vector. Furthermore, the trigonometric terms in the preceding are the same as those for $v_{z}$ in Eq. (13.4.64). This leads to the description of the phase speed as

$$
\begin{equation*}
c_{\text {phase }}=\left(\frac{\omega}{\kappa}\right)\left[1+\frac{1}{2} \beta\left(\frac{k^{3}}{\kappa^{3}}\right)\left(\frac{v_{z}}{c}\right)\right] \tag{13.4.75}
\end{equation*}
$$

If $\kappa$ were the same as $k$ and the half factor were not present, this would be the same of the phase speed of a nonlinear planar wave.

An important corollary of this analysis is the observation that rays leaving a specific point on the boundary at different instants do not coincide. This behavior is unlike the properties of rays according to the linear version of geometrical acoustics in which the rays are time invariant. Let us consider the rays that depart from a specific location $x$ on the plate's surface. Because the plate displacement is $O(\varepsilon)$, this surface is essentially $z=0$. Thus, $\theta / \omega$ is the instant when the signal departed
from the plate. According to Eq. (13.4.64), $v_{x}=0$ for $\theta=0, \pi, 2 \pi, \ldots$. It follows from Eq. (13.4.72) that the rays for these values of $\theta$ are parallel to the $z$-axis, as are the rays obtained from linear theory. The maximum positive value of $v_{x}$ corresponds to $\theta=\pi / 2,5 \pi / 2, \ldots$. Equation (13.4.72) indicates that the rays that depart from location $x$ on the plate at these instants undergo the greatest rotation to the right of the $z$-axis. Conversely, $\theta=3 \pi / 2,7 \pi / 2, \ldots$ correspond to the largest negative $v_{x}$. The rays departing at these instants undergo the greatest rotation to the left of the $z$-axis. Thus, within a period the rays rotate from one side to the other as time elapses.

An analogy explains the difference between this description of a ray and the nature of lines of constant $\eta$ in Fig. 13.23. Consider the stream of water that emerges from a hose. Suppose the nozzle is rotated about the vertical axis in an oscillatory manner. Any particle of water follows a straight line in the horizontal plane because the only force acting on that particle is gravity. (Obviously, this is not true if there is a substantial wind.) This straight line is the ray. Further suppose that we take a photograph looking down on the horizontal plane. The stream in such a photograph would be situated along a line that resembles the oscillatory constant $\eta$ line in Fig. 13.23. Both views: the path followed by a particle and the instantaneous location of signals, are meaningful, but the ray picture might be more revealing.

Another corollary of the analysis is the observation that rays for a specific phase that depart from the boundary at the same time, but different $x$, are not parallel. The maximum plate displacement $|w|$ at any instant occurs at locations where $|\sin (\pi x / L)|=1$. These are the midpoints between the locations where the plate is supported. Because $\eta=x$ at the boundary, these locations correspond to $\cos (\pi \eta / L)=0$. The deviation angle for rays leaving these locations is $\psi=0$ regardless of $\theta$. In contrast, at the locations where the plate is supported, $w=0$, which means that $\sin (\pi x / L)=\sin (\pi \eta / L)=0$. Hence, $|\cos (\pi \eta / L)|=1$ for these rays, so the fan of rays that radiate from these points is the widest.

A description of rays sometimes is accompanied by a depiction of wavefronts. Such a picture allows us to monitor the movement of the signal along its ray. We have already constructed a wavefront by fixing $\xi$ and $t$ and allowing $\eta$ to be a parameter for the coordinate transformation. Here, we wish to construct wavefronts for a specified $\theta$ at many instants. A systematic method for doing so is to treat Eqs. (13.4.70) and (13.4.71) as a parametric description of a coordinate transformation with $\theta$ and $t$ set, and $\eta$ considered to be a free parameter. Figure 13.24 displays these properties at four instants separated by a quarter period. The response is periodic in the $x$ direction, so only one interval needs to be described.)

If $\theta$ is a multiple of $\pi$, the particle velocity in the $x$ direction is zero. Consequently, the rays departing from $z=0$ at any $x$ are perpendicular to the boundary. These values of $\theta$ also correspond to $v_{z}$ having its maximum magnitude at a specified $\eta$, so the wavefronts for these phases undergo the most distortion as they propagate. The situation for $\theta$ being an odd multiple of $\pi / 2$ is opposite. These phases correspond to the magnitude of $v_{x}$ being its maximum, so the rays departing from $z=0$ at any $x$ is at the extreme angle relative to the $z$-axis. Concurrently, these phases give $v_{z}=0$, so these wavefronts remain straight as the signal propagates. Intermediate phases feature rays that are rotated less than the extreme, and wavefronts that undergo less than maximum distortion.

A consequence of the rays being rotated by different amounts is that they will intersect. This is evidenced by the region of concentration along $x=-L / 2$ for $\theta=\pi / 2$ and along $x=L / 2$ for $\theta=3 \pi / 2$. The $z$ value at which this condition occurs seems to be close to the value in Fig. 13.24 at which the coordinate straining transformation becomes multivalued. To interrogate the behavior of the rays, Fig. 13.25 zooms in on the region along $x=-L / 2$ for $\theta=\pi / 2$. This evaluation extends to a greater distance from the plate, and the density of rays close $x=0$ has been increased. Aside from shifting the $x$ values, the same picture would be obtained around the line $x=0.5 L$ for $\theta=3 \pi / 2$.


Fig. 13.24 Rays of the nonlinear signal radiated by a vibrating flat plate. The wavefronts are depicted at intervals of one-quarter period. $\varepsilon=0.005, k L=3.628$

Fig. 13.25 Zoomed view of a caustic formed by the rays of the nonlinear signal radiated by a vibrating flat plate, $\varepsilon=0.005, k L=3.628$


The enlarged figure shows that multiple rays at large angles do not have a common intersection, which means that the convergence of rays is not a focus. Rather, it is a caustic that results from the close intersection of rays at small angles relative to the line $x=-0.5 L$. This feature marks shock formation. Let us consider a range of $\theta$ values. For $\theta=0$, all rays are parallel, so the caustic is infinitely distant. As $\theta$ is increased, the caustic occurs closer to the plate somewhere along $x=-L / 2$. The caustic is closest to the plate when $\theta=\pi / 2$. Beyond that value, the caustic occurs at increasing distance, eventually ceasing to exist for $\theta=\pi$. Further increase of $\theta$ from $\pi$ to $2 \pi$ repeats this behavior along $x=0.5 L$

We can derive an expression for the distance to the caustic. The derivation is based on Fig. 13.26, which depicts a ray that intersects the line $x=-0.5 L$. The ray departs

Fig. 13.26 Intersection of a
ray departing from the plate close to the ray departing from $x=-L / 2 ; \varepsilon=0.005$, $k L=3.628, \theta=\pi / 2$

from the plate at $x=\eta$. A caustic on $x=-L / 2$ is generated by rays for which $\eta<0$, so the horizontal distance between the launch point and $x=-0.5 L$ is $-\eta-L / 2$. This is one side of a right triangle, with the opposite angle being $\tan ^{-1}(d x / d z)_{\text {ray }}$, which is given by Eq. (13.4.72). The other side of the right triangle is the distance $z_{\text {int }}$ to the intersection of the ray with $x=-L / 2$. From trigonometry, we have

$$
\begin{equation*}
z_{\mathrm{int}}=\frac{-\eta-L / 2}{\tan \left(\tan ^{-1}(d x / d z)_{\mathrm{ray}}\right)}=\frac{\eta+L / 2}{\varepsilon \beta \frac{k L}{2 \pi}\left(\frac{k^{2}}{\kappa^{2}}\right) \sin (\theta) \cos \left(\frac{\pi \eta}{L}\right)} \tag{13.4.76}
\end{equation*}
$$

The caustic occurs because the rays at small angles converge, and small angles correspond to $\eta$ being close to $-L / 2$. The arrete, which is the minimum distance to the caustic, is the limit as the angle goes to zero, that is, as $\eta \rightarrow-L / 2$. Thus

$$
\begin{equation*}
z_{\text {arrete }}=\lim _{\eta \rightarrow-\pi / 2}\left(z_{\text {int }}\right)=\frac{1}{\varepsilon \beta \frac{k}{2}\left(\frac{k^{2}}{\kappa^{2}}\right) \sin (\theta)} \tag{13.4.77}
\end{equation*}
$$

The shock formation distance is the minimum distance to the caustic at any instant, which corresponds to $\sin (\theta)=1$,

$$
\begin{equation*}
z_{\text {shock }}=\min \left(z_{\text {arrete }}\right)=\frac{2}{\varepsilon \beta k}\left(\frac{\kappa}{k}\right)^{2} \tag{13.4.78}
\end{equation*}
$$

For comparison, the plane wave shock formation distance for a harmonic excitation is $1 /(\beta \varepsilon k)$. For the parameters used to construct Fig. 13.26, the above equation gives $z_{\text {sh }} /(2 \pi / \kappa)=6.631$, which is slightly less than the distance at which two rays in Fig. 13.25 intersect.

The motivation for the analyses of the coordinate straining grid, as well as of rays and wavefronts, is to understand the processes that affect how a signal propagates. The main concepts are the existence of the self-refraction phenomenon, as well as the altered behavior of rays as properties that depend on the phase of the signal that is transported along them. These occur concurrently with amplitude dispersion, which is the process that is responsible for distortion of plane waves. However, these observations are significant, and in some respects fascinating, we also are interested in the nature of waveforms. Evaluation of the signal at specified $x, z$, and $t$ requires a quantitative solution of the coordinate straining transformation. The next example addresses the procedure for doing so, as well as the effects of the nonlinear processes.

EXAMPLE 13.8 A plate of large extent forms the surface of a large body of water. The plate is periodically supported at intervals of $L=1.5 \mathrm{~m}$. It is mechanically driven such that its displacement amplitude is $w=$ $W \cos (\omega t) \sin (\pi x / L)$. The maximum displacement is 0.5 mm , the frequency is 2500 Hz , and it may be assumed that the water remains in contact with the plate throughout the vibration. Determine the steady-state waveforms of pressure at $x=0, L / 4$, and $L / 2$ at a depth that is $95 \%$ of the shock formation distance. Also determine the particle velocity waveforms at those locations.

## Significance

Thus far, we have only examined how the rays and waveforms are affected by selfrefraction. Here, we will see its effects on waveforms. In doing so, we will see how to solve a coordinate transformation that features more than one position coordinate.

## Solution

The displacement in Eq. (13.4.40) is $w=(\varepsilon c / \omega) \sin (\omega t) \sin (\pi x / L)$. A time shift of $\pi /(2 \omega)$ will convert this expression to the given displacement. Accordingly, the derived formulas for the coordinate straining, as well as for the pressure and velocity components, are converted by setting $\varepsilon=W k$ and replacing $t$ with $t+\pi /(2 \omega)$. It is convenient to work with a functional representation of the coordinate transformation. Upon introduction of the time shift, Eq. (13.4.63) becomes

$$
\begin{equation*}
x=f_{x}(\eta, \xi, t, z), \quad z=f_{z}(\eta, \xi, t, z) \tag{1}
\end{equation*}
$$

where the functions are

$$
\begin{align*}
& f_{x}(\eta, \xi, z, t)=\eta-\frac{1}{2}(W k) \beta z \frac{k L}{\pi}\left(\frac{k}{\kappa}\right)^{2} \cos (\omega t-\kappa \xi) \cos \left(\frac{\pi \eta}{L}\right)  \tag{2}\\
& f_{z}(\eta, \xi, x, z, t)=\xi-\frac{1}{2}(W k) \beta z\left(\frac{k}{\kappa}\right)^{3} \sin (\omega t-\kappa \xi) \sin \left(\frac{\pi \eta}{L}\right)
\end{align*}
$$

Correspondingly, Eq. (13.4.64) becomes

$$
\begin{gather*}
\frac{v_{x}}{c}=-(W k) \frac{\pi}{\kappa L} \cos (\omega t-\kappa \xi) \cos \left(\frac{\pi \eta}{L}\right)+O\left(\varepsilon^{2}\right) \\
\frac{v_{z}}{c}=\frac{\kappa}{k}\left(\frac{p}{\rho_{0} c^{2}}\right)=-(W k) \sin (\omega t-\kappa \xi) \sin \left(\frac{\pi \eta}{L}\right)+O\left(\varepsilon^{2}\right) \tag{3}
\end{gather*}
$$

It is desired to determine waveforms at specified points. This requires that we solve the coordinate transformation, Eq. (13.4.63), for the values of $\eta$ and $\xi$ corresponding to specified values of $x, z$, and $t$. Repeating such a computation for a sequence of $t$ values will yield one period of a waveform, which can be periodically replicated without requiring further solutions.

There are several numerical algorithms that may be used to solve two simultaneous nonlinear equations. We shall use the extended version of Newton's method. Suppose $\eta^{(n)}$ and $\xi^{(n)}$ are estimates for the values of $\eta$ and $\xi$ that satisfy Eq. (1) at specified $x, z$, and $t$. We seek to improve that estimate, such that $\eta^{(n+1)}=\eta^{(n)}+\Delta \eta$ and $\xi^{(n+1)}=\xi^{(n)}+\Delta \xi$ satisfy both transformations, that is,

$$
f_{x}\left(\eta^{(n)}+\Delta \eta, \xi^{(n)}+\Delta \xi, t, z\right)=x, f_{z}\left(\eta^{(n)}+\Delta \eta, \xi^{(n)}+\Delta \xi, t, z\right)=z
$$

This constitutes two simultaneous nonlinear equations for $\Delta \eta$ and $\Delta \xi$. Let us expand each function in a first-order Taylor series relative to the values obtained with the prior estimates. Then the equations to solve are linear in the increments,

$$
\begin{aligned}
& \left(f_{x}\right)^{(n)}+\left(\frac{\partial f_{x}}{\partial \eta}\right)^{(n)} \Delta \eta+\left(\frac{\partial f_{x}}{\partial \xi}\right)^{(n)} \Delta \xi=x \\
& \left(f_{z}\right)^{(n)}+\left(\frac{\partial f_{z}}{\partial \eta}\right)^{(n)} \Delta \eta+\left(\frac{\partial f_{z}}{\partial \xi}\right)^{(n)} \Delta \xi=z
\end{aligned}
$$

where the superscript for the functions and their derivatives means that the quantity is evaluated for the variables at the $n$th estimate. The matrix representation of these equations is

$$
\left[J\left(f_{x}, f_{z}\right)^{(n)}\right]\left\{\begin{array}{l}
\Delta \eta \\
\Delta \eta
\end{array}\right\}=\left\{\begin{array}{l}
x-\left(f_{x}\right)^{(n)} \\
z-\left(f_{z}\right)^{(n)}
\end{array}\right\}
$$

where $\left[J\left(F_{x}, F_{z}\right)^{(n)}\right]$ is the Jacobian of the transformation evaluated at the $n$th estimate,

$$
\left[J\left(f_{x}, f_{z}\right)^{(n)}\right]=\left[\begin{array}{ll}
\left(\frac{\partial f_{x}}{\partial \eta}\right)^{(n)} & \left(\frac{\partial f_{x}}{\partial \xi}\right)^{(n)} \\
\left(\frac{\partial f_{z}}{\partial \eta}\right)^{(n)} & \left(\frac{\partial f_{z}}{\partial \xi}\right)^{(n)}
\end{array}\right]
$$

Since $\Delta \eta=\eta^{(n+1)}-\eta^{(n)}$ and $\Delta \xi=\xi^{(n+1)}-\xi^{(n)}$, the improved estimates of the strained coordinates are given by

$$
\left\{\begin{array}{l}
\eta^{(n+1)} \\
\xi^{(n+1)}
\end{array}\right\}=\left\{\begin{array}{l}
\eta^{(n)} \\
\xi^{(n)}
\end{array}\right\}+\left[J\left(f_{x}, f_{z}\right)^{(n)}\right]^{-1}\left\{\begin{array}{l}
x-\left(f_{x}\right)^{(n)} \\
z-\left(f_{z}\right)^{(n)}
\end{array}\right\}
$$

The iterative procedure requires an initial estimate. All of the results that follow were obtained by setting $\eta^{(0)}=x$ and $\xi^{(0)}=z$. The procedure terminates when a convergence criterion has been met. The one used here is based on the observation that $\left|f_{x}-x\right|$ should be much less than $\varepsilon x$ and $\left|f_{z}-z\right|$ should be much less than $\varepsilon z$. Both conditions will be met if the iteration is halted when $\left(\left(f_{x}\right)^{(n)}-x\right)^{2}+\left(\left(f_{z}\right)^{(n)}-x\right)^{2}<$ $\varepsilon^{4}\left(x^{2}+z^{2}\right)$. The iterative algorithm is used to compute the waveforms of pressure and particle velocity at each of the specified locations. To do so the procedure is carried out in a loop of $t$ values covering one period. After the values of $\eta$ and $\xi$ have been determined at each instant, the corresponding values of $p, v_{x}$, and $v_{z}$ are determined from Eq. (13.4.64). All response variables have a period of $2 \pi / \omega$, so the waveforms are replicated for another period in order to see the pattern more clearly.

The parameters for the evaluation are $c=1480 \mathrm{~m} / \mathrm{s}, \beta=3.5$ corresponding to $B / A=5, \omega=5000 \pi \mathrm{rad} / \mathrm{s}, k=10.613 \mathrm{~m}^{-1}, \kappa=10.405 \mathrm{~m}^{-1}$, and $\varepsilon=W k=$ $5.3067\left(10^{-3}\right)$. The wavelength in the $z$ direction is $2 \pi / \kappa=0.6039 \mathrm{~m}$. The shock formation distance is $x_{\text {shock }}=9.705 \mathrm{~m}$, so this is a fairly strong signal. The waveforms that are computed at $x=0,0.375$, and 0.75 , with $z=9.263 \mathrm{~m}$ are shown in Fig. 1. They are best understood if they are compared to their linear analog. Linear theory recovered by setting $\varepsilon=0$ in Eq. (2),

$$
\begin{gather*}
\frac{v_{x}}{c}=-(W k) \frac{\pi}{\kappa L} \cos (\omega t-\kappa z) \cos \left(\frac{\pi x}{L}\right) \\
\frac{v_{z}}{c}=\frac{\kappa}{k}\left(\frac{p}{\rho_{0} c^{2}}\right)=-(W k) \sin (\omega t-\kappa z) \sin \left(\frac{\pi x}{L}\right)+O\left(\varepsilon^{2}\right) \tag{5}
\end{gather*}
$$

Thus, in linear theory $p$ should be zero along $x=0$ and $v_{x}$ should exhibit its maximum excursion along that line, whereas the amplitude of $p$ should be maximum and $v_{x}$ should be zero along $x=L / 2$. The situation at $x=L / 4$ should be intermediate to these extremes.


Figure 1.
Because $v_{x}=0$ along $x=L / 2$, the ray leaving the plate at that location does not rotate. The pressure signal evolves along a ray in a manner similar to planar wave. Hence, we see the same distortion effect along this line, with positive pressures arriving earlier, and negative pressures arriving later. Along $x=0$, the pressure is nonzero because different rays pass through a point at $(0, z)$ at different instants as a result of the self-refraction effect. Rays on which $v_{x}>0$ arrive from $x<0$ for half a period, and rays with $v_{x}<0$ arrive from $x>0$ for the other period. The variation of pressure along both sets of rays is similar, so the pressure undergoes a double oscillation in one period. Hence, the pressure at that location has half of the overall period.

The waveforms for $v_{x}$ are interesting in that they are symmetric, rather than leaning over due to advancement and retardation. This symmetry occurs because $v_{x}$ in a ray coming from the plate at $x>0$ is merely $180^{\circ}$ out-of-phase from $v_{x}$ on a ray arriving from $x<0$. If we were to examine $v_{x}$ as a function of $t$ at $x=L$, we would see the inverted version of $v_{x}$ at $x=0$ because all variables are shifted there by $180^{\circ}$.

### 13.5 Further Studies

Nonlinear phenomena have been the subject of a vast body of research. The topics covered here provide the primary tools and concepts required to go further. A historical perspective is available in a comprehensive treatise by Hamilton and Blackstock, ${ }^{22}$ which is an excellent starting point for further exploration. The purpose of the discussion here is to make the reader aware of some interesting phenomena and applications without delving into the technical details.

[^81]- Radiation pressure and acoustic levitation

Our studies have concentrated on the growth of nonlinear effects as a wave propagates. The consequence was that a temporally periodic excitation with zero mean value was found to lead to a pressure field that has similar properties. This is not quite true. In actuality, there is a small mean value for the pressure at a point. This field may be traced back to the nonlinear terms that do not result in cumulative growth of distortion. For example, consider the nonlinear momentum equation, Eq. (13.3.3). It contains the term $\nabla \phi \cdot \nabla \phi$. If $\nabla \phi$ is harmonic, then $\nabla \phi \cdot \nabla \phi$ consists of a mean value and a second harmonic of equal magnitude.

The mean value of a pressure field is referred to as the radiation pressure. This leads to a radiation force, which is the mean value for the resultant force exerted on a body immersed in the pressure field. Although this concept might appear to be simple, the analysis of the concept raises some difficult issues. One concern is the obvious fact that placing a body in the field modifies the field due to scattering. That is, what is the radiation pressure distribution on the surface of a body that scatters the incident signal? Another question that arises pertains to the nature of this body. Is it fixed in space, or is it free to vibrate? Because the radiation pressure is small compared to the pressure amplitude, small differences are important, which is why the question of whether the body is restrained is an essential aspect.

One of the most interesting applications of radiation pressure is for acoustic levitation. As was noted, the radiation force is small. Using it to support a body at the earth's surface would require that the radiation force equals the body's weight, which would require an incredibly high acoustic pressure. However, for a body in orbit, for example, in the International Space Station, the weight is the centripetal force required to maintain the body in orbit. From the reference frame of the Space Station, the body is weightless and free to drift unless it is restrained. The radiation force has been used to restrain a body without contacting it. This is achieved by setting up a resonant field in a closed chamber. An unconstrained body then will drift to a pressure node. This concept is useful for any process in which contact with a container will lead to contamination or distortion of an object.

## - Sound beams

Suppose you wish to send an acoustical message to listeners at a long range but in a limited direction. We select a loudspeaker or hydrophone for this purpose, embed it in a baffle, and set the frequency sufficiently high to obtain the desired directivity. To overcome amplitude reductions due to spherical spreading and dissipation, we drive the transducer harder. Several alternatives arise. As the input is increased, the amplifier that drives the transducer might begin to distort or cut out entirely. The speaker might be overdriven, which also might cause distortion or cause the speaker to fail. The first two possibilities can be avoided with proper selection of the apparatus, but the third possibility is not avoidable. Raising the amplitude of the signal radiated by the transducer leads to the growth of nonlinear effects as the signal propagates. If the drive amplitude is sufficiently large, shock waves might form. Raising the amplitude further will merely bring the location of shocks closer to the transducer,
rather than further raising the amplitude that is received. The ultimate result might be saturation. Another possibility for an underwater signal is that the amplitude will be sufficiently large that the absolute pressure (ambient plus acoustic disturbance) is negative. The consequence of such an occurrence is creation of bubbles, which is called cavitation. This process also limits the power that can be transferred into the radiated signal. (Negative pressure is not an issue for atmospheric sound. To see why consider a plane wave for which the acoustic Mach number is $\varepsilon=|p| /\left(\rho_{0} c^{2}\right)$. A negative absolute pressure corresponds to $|p|=1 \mathrm{~atm}$, but $\rho_{0} c^{2}$ is approximately 1 $\operatorname{atm}$, and we know that $\varepsilon=1$ is only attained with an explosion. In contrast $\rho_{0} c^{2} \approx 2$ GPa for water and the ambient pressure near the surface is approximately 1 atm . Setting $|p|=1$ atm leads to $\varepsilon \approx 0.5\left(10^{-4}\right)$, which is quite attainable.)

A primary application of sound beams is a method of long-range communication. They also are at the core of some imaging concepts, such as an acoustical microscope and some biomedical ultrasonic devices. In each case, it is imperative that nonlinear effects be understood. Our study of sound beams according to linear theory required numerical methods to evaluate the general field properties. Thus, it is obvious that studying nonlinear effects is even more challenging. Many analyses, particularly the early ones were based on a regular perturbation analysis, in which the linear solution for a sound beam is used to describe the second order, that is, quadratic, nonlinear terms. The Rayleigh integral is one way of describing the linear solution, but there are others. Furthermore, to reduce the level of difficulty various simplified field equations often provided the basis for the analysis. The simplifications were based on properties of sound beams at high frequency, such as that the spatial scale of diffraction is large compared to the axial wavelength, but small compared to the Rayleigh distance.

A regular perturbation analysis is limited in its range of validity. A more complete picture of nonlinear effects was obtained by using some of the methods developed here, notably the frequency domain Fourier series approach and the singular perturbation formulation.

## - Parametric arrays

An application that relies on nonlinear effects is a concept that combats a basic aspect of the linear field of a sound beam. Specifically, if one wishes the beamwidth to be very small, then the frequency must be high, but high frequencies mean greater dissipation, thereby reducing the effective range. The concept entails imparting to the piston on the boundary a vibration that is the sum of two harmonic signals at closely spaced frequencies. The linear field, which is the first-order solution for a regular perturbation series, is the superposition of sound beams at each frequency. If both frequencies are high, both beams will be narrow, and their directivities will be quite similar because the frequencies are close. When the first-order solution is used to generate the quadratic terms in a nonlinear wave equation, approximate or exact, various combinations of harmonic signals are created. Second harmonics of each primary frequency are like the behavior of each beam if it were the only one present. Mixing of the two harmonics results in a sum frequency signal, which is close to the second harmonics of the individual beams, and a difference frequency signal. The latter part
is the one that is of interest because proximity of the two drive frequencies means that the difference frequency is very small. Furthermore, because these terms are the product of two narrow directivities, the difference frequency terms are significant only near the axis of the sound beam. The homogeneous terms of the second-order wave equation are the same as those of the first-order equation, but the quadratic terms occur on the right side of the equation as virtual sources. These difference frequency sources are confined to the region close to the axis. In combination with the fact that the propagation is predominantly axial, the consequence is that the virtual sources act like a narrow end-fire array. The creation of this array is attributable to the parameters of the primary frequency sound beams and hence the name of the difference frequency sources as a parametric array. This concept has been used in some applications, notably as a narrowly confined megaphone. It has not been used widely because it relies on nonlinear interactions, which are relatively weak, with a magnitude that is the order of the square of the acoustic Mach number.

## - Biomedical applications

The use of sound beams for ultrasonic diagnostics was mentioned earlier. Increasing the depth of penetration by raising the amplitude enhances nonlinear effects. One aspect of this is beneficial in that higher harmonics have smaller wavelengths. This serves to enhance the visual resolution. At the same time, larger amplitudes lead to greater tissue heating. This is a harmful effect as a diagnostic tool. However, it is useful as a method for destroying tissue that is harmful, such as cancer cells. One of the key requirements for understanding these phenomena is a description of how large amplitude sound waves propagate through living tissue, which is heterogeneous and dissipative through heat transport and viscosity. This is a subject of ongoing research.

Lithotripsy is a different way in that finite amplitude effects are exploited. Stones can form in several organs, such as kidneys. They can become extremely painful, in which case highly invasive surgery is the usual treatment. A lithotripter provides a nonsurgical alternative. The basic concept is to immerse the torso in a tub of water, with the walls of the tub driven as a pulse. This generates an inwardly propagating transient wave that converges on the stone. The consequence of convergence is that the pressure field applied to the stone is greatly enhanced and changes even more abruptly than the original pulse. (Recall the behavior in linear theory of an inwardly propagating spherical wave at the focus, see Sect.6.2.3.) Hence, the pressure acting on the stone is large and the pressure gradient is very large. Furthermore, the stone typically is somewhat irregular in shape and has some internal faults. The combination of these attributes is that large stresses induced within the stone might cause it to fracture into smaller pieces that might cause it to pass through the duct. Transfer from a concept to a functioning apparatus required a considerable body of research concerning focusing of nonlinear waves and their interaction with solid objects.

### 13.6 Homework Exercises

Exercise 13.1 An explosion in air produces a pressure pulse in the form of a single lobe of a sine function. The amplitude is 20 kPa and the duration is 0.1 s . Draw the characteristics that track the peak pressure, and the zero pressure values at the beginning and end of the pulse. Identify the smallest distance at which any two of these lines intersect. Sketch the waveform predicted by the Riemann solution at this location. What does that sketch indicate regarding the location at which a shock forms?

Exercise 13.2 A transducer at the near end, $x=0$, of a very long tube imparts a velocity wave form in the shape of a triangular pulse of duration $T$, given by

$$
v=\left\{\begin{array}{l}
\varepsilon c \frac{2 t-T}{T}, \quad 0<t<T \\
0 \text { otherwise }
\end{array}\right.
$$

Thus, $t=T$ is the last instant at which the transducer is moving. (a) Draw the spatial profile of this wave at $t=T$. (b) Evaluate the shock formation distance $x_{s}$ for this wave. (c) Evaluate the waveform at $x=x_{s}$.

Exercise 13.3 At $t=0$ the signal in a waveguide is observed to be a plane wave propagating in the direction of increasing $x$. Its spatial profile at that instant has the shape of an equilateral triangle,

$$
\frac{p}{\rho_{0} c^{2}}=\left\{\begin{array}{l}
\varepsilon \frac{2 x}{L}, \quad 0<x<L / 2 \\
\varepsilon \frac{2(L-x)}{L}, \quad L / 2<x<L \\
0 \text { otherwise }
\end{array}\right.
$$

Determine the shock formation distance $x_{\text {shock }}$ according to weak shock theory. Then draw the waveform at $x=x_{\text {shock }}$ and $x=x_{\text {shock }} / 2$.

Exercise 13.4 A projectile impacts the back of a wall, thereby inducing a displacement pulse $u=(4 / 3) \varepsilon c\left(t^{2} / T^{2}\right)(3 T-2 t)$ that occurs over $0 \leq t \leq T$. The parameters are $\varepsilon=0.0001$ and $T=10 \mathrm{~ms}$, and the fluid is air at standard conditions. (a) Determine the minimum distance at which a shock forms according to the weakly nonlinear approximation. (b) Use the weakly nonlinear approximation of the Riemann solution to evaluate the waveform at $x=x_{\text {shock }}$. (c) Use Eq.(13.1.17), which accounts for movement of the wall, to evaluate the waveform at $x=x_{\text {shock }}$. Estimate the overall error in the result in Part (b) that results from using a stationary, rather than moving, boundary condition.

Exercise 13.5 The velocity of a vibrating wall consists of a pulse that occurs over an interval $T$. The waveform of that pulse is a single cycle of a sine function, $v=$ $\varepsilon c \sin (2 \pi t / T)$. The fluid is air at standard conditions. To test whether the limits of the weakly nonlinear approximation set $\varepsilon=0.1, T=4 \mathrm{~ms}$. (a) Use the weakly nonlinear approximation to evaluate the waveform at the shock formation distance.
(b) Use the exact ideal gas relations to evaluate the waveform at the same location.

Exercise 13.6 If the particle velocity in a weakly nonlinear plane wave at $x=0$ is $v=\varepsilon c \sin (\omega t)$, then the shock formation distance is $x_{\text {shock }}=1 /(\beta \varepsilon k)$. Consider the case of a 1 kHz signal in water with $\varepsilon=0.0001$. (a) Use numerical methods to evaluate and graph the waveform at $x=x_{\text {shock }}$. (b) Evaluate and graph the waveform at $x=1 /(\beta \varepsilon k)$ when the input waveform (at $x=0$ ) is the negative of the waveform at $x=x_{\text {shock }}$ found in Part (a). (c) Based on the properties of the waveforms at $x=0$ and $x=1 /(\beta \varepsilon k)$, identify the minimum value of $x$ at which the inverted signal will form a shock.

Exercise 13.7 An experiment in the Tunnel de Sainte-Marie-aux-Mines in eastern France measured the waveform of a pressure pulse at 3 Km from the western entrance. With $t=0$ defined as the instant when the pulse arrived, a function fitting the measured data is $p=42(t / T)^{1 / 2}(1-t / T)[h(t)-h(t-1.25 T)] \mathrm{Pa}$, with $T=12 \mathrm{~ms}$. Because $\partial p / \partial t$ is infinite at $t=0$ and $t=1.25 T$, with $p$ increasing at both instants, the location where this waveform was measured marks shock formation. Determine the waveform at the tunnel entrance, where the pulse was generated. It may be assumed that waves only propagate from the source to the measurement point, and that the propagation is planar with negligible dissipation.

Exercise 13.8 The pressure input at $x=0$ consists of second and third harmonics, such that

$$
p=\varepsilon \rho_{0} c^{2}\left[a_{2} \sin (2 \omega t)+a_{3} \sin (3 \omega t)\right]
$$

When $a_{3}=0$, the shock forms at $x=1 /\left(2 a_{2} \varepsilon \beta k\right)$, whereas this location is $k x=$ $1 /\left(3 a_{3} \varepsilon \beta\right)$ when $a_{2}=0$, where $k=\omega / c$. Cases of interest are $a_{2}=1, a_{3}=0.1$, and $a_{2}=1, a_{3}=2 / 3$. For both cases $\varepsilon=0.002$ and $\beta=3.6$. Evaluate and graph $p /\left(\rho_{0} c^{2}\right)$ as a function of $\omega t-k x$ at $k x=0.25 /\left(\beta \varepsilon a_{2}\right)$ for each case. Compare each waveform to its linear counterpart. Identify the instants at which $\partial p / \partial t$ is greatest in each case, and examine each waveform to determine whether a shock has formed.

Exercise 13.9 The pressure input at $x=0$ consists of second and third harmonics, such that

$$
p=\varepsilon \rho_{0} c^{2}[\sin (2 \omega t)+(2 / 3) \sin (3 \omega t)]
$$

The period of this function is $\omega T=2 \pi$. The parameters are $\varepsilon=0.002$ and $\beta=3.5$. (a) Use numerical methods to analyze the waveform at a sequence of position $k x_{n}=$ $0.04 n /(\beta \varepsilon), n=0,1, \ldots, 6$. (b) Use FFT techniques to evaluate the amplitude of harmonics 1 to 5 at each position.

Exercise 13.10 Consider the situation in Exercise 13.9. Rather than using FFT analysis of computed waveforms, use Fourier series analysis of the Riemann solution to determine the amplitude of harmonics 1 to 5 at $k x_{n}=0.04 n /(\beta \varepsilon), n=0,1, \ldots, 6$.

Exercise 13.11 The pressure input at $x=0$ consists of second and third harmonics, such that

$$
p=\varepsilon \rho_{0} c^{2}[\sin (2 \omega t)+(2 / 3) \sin (3 \omega t)]
$$

The fluid is air with the attenuation described by Eq.(3.3.33). The parameters are $\omega=800 \mathrm{rad} / \mathrm{s}, \varepsilon=0.002$, and $\beta=1.2$. Solve the differential equations governing the waveform's Fourier series coefficients in order to evaluate the $x$ dependence of the lowest four harmonic amplitudes. Halt the evaluation at $k x=0.24 /(\beta \varepsilon)$. Compare the waveform at the farthest location to the result that would be obtained if dissipation were negligible.

Exercise 13.12 The pressure generated by a periodic source at $x=0$ is a combination of the ninth and tenth harmonics, specifically, $p=\rho_{0} c^{2} \varepsilon(\sin (9 \omega t)+\sin$ $(10 \omega t))$. At the origin, this signal constitutes a beat, but an important aspect of the nonlinear process is harmonic and intermodulation distortion, in which higher and lower harmonics are generated. Derive an expression for the shock formation distance, $x_{\text {shock }}$. Then, solve the differential equations in Sect.13.3.2 for $0 \leq x \leq$ $0.9 x_{\text {shock }}$. Use Eq. (3.3.33) to describe the frequency dependence of the attenuation coefficient. The system parameters are $\omega /(2 \pi)=500 \mathrm{~Hz}$ and $\varepsilon=0.002$. The fluid is water, for which $\rho_{0}=1000 \mathrm{~kg} / \mathrm{m}^{3}, c=1480 \mathrm{~m} / \mathrm{s}$, and $\beta=3.6$. Plot the amplitudes of pressure harmonics 1 to 15 as a function of $x$. Use a logarithmic scale for the amplitude values.

Exercise 13.13 The spatial profile of a wave is described as $p$ as a function of $x$ at fixed $t$. Prove that the shock formation distance is manifested in a profile as an infinite value of $\partial p / \partial x$ with $p$ decreasing.

Exercise 13.14 A piston at the end $x=0$ of a semi-infinite waveguide vibrates periodically. Its steady-state velocity is the sawtooth waveform in the sketch. The waveform at any $x$ preceding shock formation is a scalene triangle. (a) Use the graphical construction method to establish the shape of the triangle at an arbitrary $x$ preceding shock formation. (b) Carry out a Fourier analysis of the waveform in Part (a), and thereby, derive an expression for the dependence of the harmonic amplitudes on $x$. (c) Prove that the shock formation distance for this signal is $x_{\text {shock }}$ $=c T /(2 \varepsilon \beta)$. (d) Evaluate the waveform and harmonic content at $x_{n}=n x_{\text {shock }} / 4$, $n=0,1, \ldots, 4$. Which of these harmonics are enhanced and which are depleted as the signal propagates?


Exercise 13.14
Exercise 13.15 The periodic boundary velocity in Exercise 13.14 generates a steadystate plane wave in an infinitely long waveguide. (a) Determine the distance $x_{\text {old }}$ at which the wave enters the old-age stage. (b) Determine the waveform at $x>x_{\text {old }}$. From that result, derive an expression for the maximum pressure as a function of $x$.

Exercise 13.16 A pressure surge occurs at $x=0$ in a waveguide filled with water. The duration of the pulse is $0 \leq t \leq T$, and its signature in that interval is $p(x=0, t)=0.50\left(10^{6}\right)(t / T) \exp (3(1-t / T)) \mathrm{Pa}$ where $T=50 \mathrm{~ms}$. (a) Determine the shock formation distance $x_{\text {shock. }}$. (b) Evaluate the Riemann solution at evenly spaced locations $x=x_{\text {old }}, 2 x_{\text {old }}, 3 x_{\text {old }}, \ldots$ Sequentially examine each waveform to estimate the smallest distance at which old age begins. (c) Evaluate the waveform at the location estimated in Part (b) to be the beginning of old age. Does the shock formation distance obtained from the analysis indicate that the estimated distance was too large or too small?

Exercise 13.17 When the spring in a one-degree-of-freedom oscillator is weakly nonlinear, the force it exerts deviates from proportionality to its elongation. A model that is descriptive of some systems adds a force term that is proportional to the square of the displacement. The resulting equation of motion is

$$
M \ddot{y}+\mu \dot{y}+K_{1} y+K_{2} y^{2}=f(t)
$$

where $f(t)$ is the excitation force and $\mu$ is a viscosity coefficient that accounts for dissipation. In the case where $f(t)$ is a periodic function, the steady-state response must share that period. This means that $y$, as well as $f$, may be represented by Fourier series. Thus, let

$$
y=\frac{1}{2} \sum_{n=-N}^{N} Y_{n} e^{i n \omega t}, \quad Y_{(-n)}=Y_{n}^{*} f=\frac{1}{2} \sum_{n=-N}^{N} F_{n} e^{i n \omega t}, \quad F_{(-n)}=F_{n}^{*}
$$

where the harmonic $N$ for truncation is based on convergence of the series. Derive a set of algebraic equations whose solution would give the $Y_{n}$ values when the $F_{n}$ values are specified.

Exercise 13.18 If an elastic beam is fastened at its ends to support that are immobile, any displacement increases the length of the beam. This elongation is an accumulation
of the effect of displacement $w$ along the beams length, so it is represented by an integral. The effect of the elongation is to make the beam stiffer. The field equations and boundary conditions for the beam are

$$
E I \frac{\partial^{4} w}{\partial x^{4}}+\frac{1}{2} \frac{E A}{L^{2}} \int_{0}^{L}\left(\frac{\partial w}{\partial x}\right)^{2} d x+\sigma \frac{\partial^{2} w}{\partial t^{2}}=q
$$

where $E$ is Young's modulus, $A$ and $I$ are the area and the area moment of inertia of the cross section, $\sigma$ is the mass per unit length, and $q$ is the force loading per unit length of the beam. If the beam is supported at its ends by stationary pins that allow the beam to rotate freely, they are said to be simple supports. The associated boundary conditions are $w=\partial^{2} w / \partial x^{2}=0$ at $x=0$ and $x=L$. If the displacement is small relative to the span $L$, then it may be represented nondimensionally as $w=\varepsilon L \xi(x, t)$, where $|\varepsilon| \ll 1$. Write the dimensionless displacement as a regular perturbation series, $\xi=\xi_{1}+\varepsilon \xi_{2}+\cdots$. Derive the differential equation and boundary conditions governing $\xi_{1}$ and $\xi_{2}$.

Exercise 13.19 When the spring in a one-degree-of-freedom oscillator is weakly nonlinear, the force it exerts deviates from proportionality to its elongation. This effect often is modeled by adding a force term that is proportional to the cube of the displacement. The resulting equation of motion is

$$
M \ddot{x}+K_{1} x+K_{3} x^{3}=F(t)
$$

where $F(t)$ is the excitation force. The nonlinear coefficient $K_{3}$ is much less than $K_{1}$, so it may be written as $K_{3}=\varepsilon K_{1},|\varepsilon| \ll 1$. A regular perturbation series for the oscillator displacement is $x=x_{0}+\varepsilon x_{1}+\cdots$. (a) Derive the differential equations governing $x_{1}, x_{2}$, and $x_{3}$. (b) When the force varies harmonically, $F(t)=\hat{F} \cos (\omega t)$, the steady-state response also must be periodic. Find the steady-state response of the differential equations obtained in Part (a). Is there any value of $\omega$ at which the solution fails to be uniformly valid?

Exercise 13.20 Free vibration of the oscillator in Exercise 13.19 corresponds to $f(t)=0$. Consider an initial state in which $x=X_{0}$ and $\dot{x}=0$ when $t=0$. Division of the equation of motion by $M$ leads to

$$
\ddot{x}+\omega_{\mathrm{nat}}^{2}\left(x+\varepsilon x^{3}\right)=0
$$

where $\omega_{\text {nat }}=\left(K_{1} / M\right)^{1 / 2}$ is the natural frequency of the linear system. Use the regular perturbation series $x=x_{0}+\varepsilon x_{1}+\cdots$ to obtain a solution of the equation of motion that is consistent with the initial conditions. Then, apply the method of strained coordinates to the time variable in order to obtain a uniformly valid description of the response.

Exercise 13.21 The diagram is the pressure pulse measured on the surface of a sphere, $r=a$. This radius is much greater than the spatial extent of the pulse, that is, $a \gg c T$. (a) Use the graphical construction method to sketch the pressure waveform at an arbitrary radial distance. (b) Use the waveform obtained in Part (a) to derive an expression for the radial distance $r_{\text {shock }}$ at which a shock forms. (c) Determine the waveforms of the radial particle velocity at $r=a, r=r_{\text {shock }} / 2$, and $r=r_{\text {shock }}$.


Exercise 13.21.
Exercise 13.22 The pressure resulting from a huge explosion in the atmosphere is measured at $r_{1}=2 \mathrm{~km}$ from the ignition point. A function that correlates well with the measured data is $p=p_{1}(t / T)(1-5 t / T) \exp (-\mu t)$, with $p_{1}=20 \mathrm{kPa}$ and $\mu=4.54 \mathrm{~s}^{-1}$. This function applies for $t<2$ seconds, after which it may be set to zero. It is believed that the propagation is well described by the weakly nonlinear approximation beyond $r_{0}=5 \mathrm{~m}$. (Inside this distance, the signal is highly nonlinear and the water is in a nonideal state.) Plot $p$ at $r_{0}$ as a function of the retarded time $t^{\prime}=$ $t-\left(r_{0}-r_{1}\right) / c$. Without performing calculations, conjecture what the waveform would look like at a radial distance $r$ that is significantly greater than $r_{1}$ ?

Exercise 13.23 A 10 m diameter sphere will be used as the living chamber of a deepwater submersible. One of its tests will apply an internal pressure fluctuation when it is submerged. This pressure induces a radially symmetric vibration of the outer surface in which the surface velocity is $v_{r}=0$ for $t<0$ and $t>T$, and $v_{r}=\varepsilon c(1-4 / T)^{2}(5 t / T-1)$ for $0<t<T$. Parameters are $\varepsilon=0.0032$ and $T=0.94 \mathrm{~ms}$. (a) Determine the pressure waveform at the radial distance at which a shock develops. (b) Determine the radial velocity at $r=a$ and $r=r_{\text {shock }}$. Plot $v_{r} / c$ at each location as a function of the retarded time $t-(r-a) / c)$. Compare each waveform to $(r / a) p /\left(\rho_{0} c^{2}\right)$ as a function of the retarded time. What can be deduced about the distance at which the farfield begins?

Exercise 13.24 If $-\dot{F}(\tau)$ in Eq.(13.4.29) is replaced by $c^{2} V(\tau)$, the resulting expression for $(r / a) p$ in a spherical wave has the same appearance as $p=\rho_{0} c^{2} \varepsilon V(\tau)$ for a plane wave. Use this similarity to modify the general Fourier series analysis in Sect. 13.2.1 in order to determine the Fourier coefficients of a nonlinear spherical wave when the function $\dot{F}(t)$ is periodic. Specialize the general result to the case where the pressure at $x=0$ is a single harmonic.

Exercise 13.25 At a reference radial distance $r_{0}$, the waveform of a radially symmetric spherical wave is observed to be the square wave depicted below. (a) Derive an expression for the radial distance $r_{\text {old }}$ at which this wave enters old age. (b) Derive
an expression for the maximum positive pressure in this wave if the radial distance exceeds $r_{\text {old }}$.


Exercise 13.25.

## Appendix A Curvilinear Coordinates

## A. 1 Spherical Coordinates

When $p$ is known in terms of the spherical coordinates of the location at which is observed, the most expedient evaluation of $\nabla p$ uses a form in terms of the radial, polar, and azimuthal components. This form is more complicated in appearance than merely the derivative of a component with respect to a coordinate because incrementing the angular variables alters the direction of the unit vectors. There are several approaches by which these effects may be described in a derivation of the gradient $\nabla p$ and the Laplacian $\nabla^{2} p$ in spherical coordinates. We shall employ a direct approach based on the coordinate transformation to derive $\nabla p$, and then switch to a slightly more subtle approach based on the properties of the spherical coordinate unit vectors to derive $\nabla^{2} p$.

## A.1.1 Transformations

The derivation of a gradient in spherical coordinates begins by considering the pressure to be a function of those coordinates. In turn, the spherical coordinates are treated as functions of the Cartesian coordinates. Doing so calls for the chain rule for partial differentiation. It is helpful to have the transformations at hand. The coordinates are defined in Fig. A.1.

The Cartesian coordinates corresponding to a given a set of spherical coordinates are

$$
\begin{gather*}
x=r \sin \psi \cos \theta \\
y=r \sin \psi \sin \theta  \tag{A.1.1}\\
z=r \cos \psi
\end{gather*}
$$

The inverse transformation gives the spherical values for a given set of Cartesian coordinates. It is

Fig. A. 1 Spherical coordinates and associated unit vectors


$$
\begin{gather*}
r=\left(x^{2}+y^{2}+z^{2}\right)^{1 / 2} \\
\psi=\cos ^{-1}\left(\frac{z}{\left(x^{2}+y^{2}+z^{2}\right)^{1 / 2}}\right)  \tag{A.1.2}\\
\theta=\tan ^{-1}\left(\frac{y}{x}\right)=\sin ^{-1}\left(\frac{y}{\left(x^{2}+y^{2}\right)^{1 / 2}}\right)=\cos ^{-1}\left(\frac{x}{\left(x^{2}+y^{2}\right)^{1 / 2}}\right)
\end{gather*}
$$

We also will have need for the definition of spherical unit vectors. Each vector is the change of the position vector from the origin when the corresponding coordinate is increased by a unit value, with the others held constant. The result of doing so is

$$
\begin{gather*}
\bar{e}_{r}=\sin \psi \cos \theta \bar{e}_{x}+\sin \psi \sin \theta \bar{e}_{y}+\cos \psi \bar{e}_{z} \\
\bar{e}_{\psi}=\cos \psi \cos \theta \bar{e}_{x}+\cos \psi \sin \theta \bar{e}_{y}-\sin \psi \bar{e}_{z}  \tag{A.1.3}\\
\bar{e}_{\theta}=-\sin \theta \bar{e}_{x}+\cos \theta \bar{e}_{y}
\end{gather*}
$$

## A.1.2 Gradient

The Cartesian components of a gradient are the derivative with respect to the corresponding coordinate. The chain rule is needed to evaluate these derivatives when $p$ is known in terms of $r, \psi$, and $\theta$, so we have

$$
\begin{align*}
& \frac{\partial p}{\partial x}=\frac{\partial p}{\partial r} \frac{\partial r}{\partial x}+\frac{\partial p}{\partial \psi} \frac{\partial \psi}{\partial x}+\frac{\partial p}{\partial \theta} \frac{\partial \theta}{\partial x} \\
& \frac{\partial p}{\partial y}=\frac{\partial p}{\partial r} \frac{\partial r}{\partial y}+\frac{\partial p}{\partial \psi} \frac{\partial \psi}{\partial y}+\frac{\partial p}{\partial \theta} \frac{\partial \theta}{\partial y}  \tag{A.1.4}\\
& \frac{\partial p}{\partial z}=\frac{\partial p}{\partial r} \frac{\partial r}{\partial z}+\frac{\partial p}{\partial \psi} \frac{\partial \psi}{\partial z}+\frac{\partial p}{\partial \theta} \frac{\partial \theta}{\partial z}
\end{align*}
$$

The transformation from Cartesian to spherical coordinates is used to fill in the derivatives appearing in these equations. Implicit differentiation is useful for the angles. The derivatives of Eq. (A.1.2) with respect to $x$ proceed as follows:

$$
\begin{align*}
\frac{\partial r}{\partial x} & =\frac{\partial}{\partial x}\left(x^{2}+y^{2}+z^{2}\right)^{1 / 2}=\frac{x}{\left(x^{2}+y^{2}+z^{2}\right)^{1 / 2}} \\
\frac{\partial}{\partial x}(\cos \psi) & \equiv-\sin \psi \frac{\partial \psi}{\partial x}=\frac{\partial}{\partial x}\left(\frac{z}{r}\right)=\frac{\partial}{\partial x}\left(\frac{z}{\left(x^{2}+y^{2}+z^{2}\right)^{1 / 2}}\right) \\
& =-\frac{x z}{\left(x^{2}+y^{2}+z^{2}\right)^{3 / 2}} \\
\frac{\partial}{\partial x}(\tan \theta) & \equiv \frac{1}{(\cos \theta)^{2}} \frac{\partial \theta}{\partial x}=\frac{\partial}{\partial x}\left(\frac{y}{x}\right)=-\frac{y}{x^{2}} \tag{A.1.5}
\end{align*}
$$

Derivatives with respect to $y$ and $z$ are obtained by similar operations.
We seek a description of $\nabla p$ in which only variables associated with spherical coordinates appear. The transformations from Cartesian to spherical coordinates, along with $x^{2}+y^{2}+z^{2}=r^{2}$, are used to eliminate $x, y$, and $z$ from the preceding. Doing so leads to

$$
\begin{array}{cll}
\frac{\partial r}{\partial x}=\sin \psi \cos \theta, & \frac{\partial r}{\partial y}=\sin \psi \sin \theta, & \frac{\partial r}{\partial z}=\cos \psi \\
\frac{\partial \psi}{\partial x}=\frac{\cos \psi \cos \theta}{r}, \quad \frac{\partial \psi}{\partial y}=\frac{\cos \psi \sin \theta}{r}, & \frac{\partial \psi}{\partial z}=-\frac{\sin \psi}{r}  \tag{A.1.6}\\
\frac{\partial \theta}{\partial x}=-\frac{\sin \theta}{r \sin \psi}, \quad \frac{\partial \theta}{\partial y}=\frac{\cos \theta}{r \sin \psi}, & \frac{\partial \theta}{\partial z}=0
\end{array}
$$

These expressions are substituted into Eq. (A.1.4), which then are used to form the gradient in Cartesian components. This step yields

$$
\begin{align*}
\nabla p & =\left(\frac{\partial p}{\partial r} \sin \psi \cos \theta+\frac{1}{r} \frac{\partial p}{\partial \psi} \cos \psi \cos \theta-\frac{1}{r} \frac{\partial p}{\partial \theta} \frac{\sin \theta}{\sin \psi}\right) \bar{e}_{x} \\
& +\left(\frac{\partial p}{\partial r} \sin \psi \sin \theta+\frac{1}{r} \frac{\partial p}{\partial \psi} \cos \psi \sin \theta+\frac{1}{r} \frac{\partial p}{\partial \theta} \frac{\cos \theta}{\sin \psi}\right) \bar{e}_{y}  \tag{A.1.7}\\
& +\left(\frac{\partial p}{\partial r} \cos \psi-\frac{1}{r} \frac{\partial p}{\partial \psi} \sin \psi\right) \bar{e}_{z}
\end{align*}
$$

The last step is to convert this expression into spherical coordinate components. By definition, a component is the projection of a vector onto a coordinate axis, so we may write

$$
\begin{equation*}
\nabla p \equiv\left(\nabla p \cdot \bar{e}_{r}\right) \bar{e}_{r}+\left(\nabla p \cdot \bar{e}_{\psi}\right) \bar{e}_{\psi}+\left(\nabla p \cdot \bar{e}_{\theta}\right) \bar{e}_{\theta} \tag{A.1.8}
\end{equation*}
$$

Substitution of the unit vectors in Eq. (A.1.3) and the gradient on Eq. (A.1.7) ultimately reduces to

$$
\begin{equation*}
\nabla p=\frac{\partial p}{\partial r} \bar{e}_{r}+\frac{1}{r} \frac{\partial p}{\partial \psi} \bar{e}_{\psi}+\frac{1}{r \sin \psi} \frac{\partial p}{\partial \theta} \bar{e}_{\theta} \tag{A.1.9}
\end{equation*}
$$

One device for remembering this expression is that $d r$ is the displacement when $r$ is incremented infinitesimally, $r d \psi$ is the meridional displacement for an infinitesimal increment of $\psi$, and $(r \sin \psi) d \theta$ is the azimuthal displacement when $\theta$ is incremented.

## A.1.3 Laplacian

The final form of the gradient in spherical coordinates is understandable as changes in the field when the position is shifted infinitesimally in each coordinate direction. The Laplacian is less intuitive and somewhat more difficult to derive. One approach for deriving $\nabla^{2} p$ parallels that used to describe $\nabla p$.Specifically, it applies the chain rule to convert $(\partial / \partial x)(\partial p / \partial x),(\partial / \partial y)(\partial p / \partial y)$, and $(\partial / \partial z)(\partial p / \partial z)$ to derivatives with respect to $r, \psi$, and $\theta$. The derivation that follows proceeds differently. It uses Eq. (A.1.9) to describe both terms in $\nabla^{2} p \equiv \nabla \cdot \nabla p$, so that

$$
\begin{align*}
\nabla^{2} p= & \left(\frac{\partial}{\partial r} \bar{e}_{r}+\frac{1}{r} \frac{\partial}{\partial \psi} \bar{e}_{\psi}+\frac{1}{r \sin \psi} \frac{\partial}{\partial \theta} \bar{e}_{\theta}\right) \cdot\left(\frac{\partial p}{\partial r} \bar{e}_{r}\right. \\
& \left.+\frac{1}{r} \frac{\partial p}{\partial \psi} \bar{e}_{\psi}+\frac{1}{r \sin \psi} \frac{\partial p}{\partial \theta} \bar{e}_{\theta}\right) \tag{A.1.10}
\end{align*}
$$

Neither the components of $\nabla p$ nor the unit vectors are constant. For this reason, the derivatives must be evaluated prior to taking dot products. In other words,

$$
\begin{align*}
\nabla^{2} p= & \bar{e}_{r} \cdot \frac{\partial}{\partial r}\left(\frac{\partial p}{\partial r} \bar{e}_{r}+\frac{1}{r} \frac{\partial p}{\partial \psi} \bar{e}_{\psi}+\frac{1}{r \sin \psi} \frac{\partial p}{\partial \theta} \bar{e}_{\theta}\right) \\
& +\frac{1}{r} \bar{e}_{\psi} \cdot \frac{\partial}{\partial \psi}\left(\frac{\partial p}{\partial r} \bar{e}_{r}+\frac{1}{r} \frac{\partial p}{\partial \psi} \bar{e}_{\psi}+\frac{1}{r \sin \psi} \frac{\partial p}{\partial \theta} \bar{e}_{\theta}\right)  \tag{A.1.11}\\
& +\frac{1}{r \sin \psi} \bar{e}_{\theta} \cdot \frac{\partial}{\partial \theta}\left(\frac{\partial p}{\partial r} \bar{e}_{r}+\frac{1}{r} \frac{\partial p}{\partial \psi} \bar{e}_{\psi}+\frac{1}{r \sin \psi} \frac{\partial p}{\partial \theta} \bar{e}_{\theta}\right)
\end{align*}
$$

Further progress entails evaluation of the derivatives of the unit vectors with respect to each spherical coordinate. These operations are applied to the unit vectors in Eq. (A.1.3). For example,

$$
\begin{equation*}
\frac{\partial \bar{e}_{\psi}}{\partial \psi}=-\sin \psi \cos \theta \bar{e}_{x}-\sin \psi \sin \theta \bar{e}_{y}-\cos \psi \bar{e}_{z} \tag{A.1.12}
\end{equation*}
$$

It will be easier to evaluate the dot products in Eq. (A.1.11) if the unit vector derivatives are expressed in terms of $\bar{e}_{r}, \bar{e}_{\psi}$, and $\bar{e}_{\theta}$. Toward that end we form the dot product of each unit vector with the derivative of each unit vector. For the preceding description of $\partial \bar{e}_{\psi} / \partial \psi$, these operations yield

$$
\begin{align*}
& \bar{e}_{r} \cdot \frac{\partial \bar{e}_{\psi}}{\partial \psi}=\left(\sin \psi \cos \theta \bar{e}_{x}+\sin \psi \sin \theta \bar{e}_{y}+\cos \psi \bar{e}_{z}\right) \cdot \frac{\partial \bar{e}_{\theta}}{\partial \psi}=-1 \\
& \bar{e}_{\psi} \cdot \frac{\partial \bar{e}_{\psi}}{\partial \psi}=\left(-\cos \psi \cos \theta \bar{e}_{x}-\cos \psi \sin \theta \bar{e}_{y}-\sin \psi \bar{e}_{z}\right) \cdot \frac{\partial \bar{e}_{\psi}}{\partial \psi}=0  \tag{A.1.13}\\
& \bar{e}_{\theta} \cdot \frac{\partial \bar{e}_{\psi}}{\partial \psi}=\left(-\sin \theta \bar{e}_{x}+\cos \theta \bar{e}_{y}\right) \cdot \frac{\partial \bar{e}_{\psi}}{\partial \psi}=0
\end{align*}
$$

These components are used to construct the vector representation of $\partial \bar{e}_{\theta} / \partial \theta$ according to

$$
\begin{equation*}
\frac{\partial \bar{e}_{\psi}}{\partial \psi}=\left(\bar{e}_{r} \cdot \frac{\partial \bar{e}_{\psi}}{\partial \psi}\right) \bar{e}_{r}+\left(\bar{e}_{\psi} \cdot \frac{\partial \bar{e}_{\psi}}{\partial \psi}\right) \bar{e}_{\psi}+\left(\bar{e}_{\theta} \cdot \frac{\partial \bar{e}_{\psi}}{\partial \psi}\right) \bar{e}_{\theta}=-\bar{e}_{r} \tag{A.1.14}
\end{equation*}
$$

The other derivatives are found by a similar process. The full set of results is

$$
\begin{gather*}
\frac{\partial \bar{e}_{r}}{\partial r}=\overline{0}, \quad \frac{\partial \bar{e}_{\psi}}{\partial r}=\overline{0}, \quad \frac{\partial \bar{e}_{\theta}}{\partial r}=\overline{0} \\
\frac{\partial \bar{e}_{r}}{\partial \psi}=\bar{e}_{\psi}, \quad \frac{\partial \bar{e}_{\psi}}{\partial \psi}=-\bar{e}_{r}, \quad \frac{\partial \bar{e}_{\theta}}{\partial \psi}=\overline{0}  \tag{A.1.15}\\
\frac{\partial \bar{e}_{r}}{\partial \theta}=\sin \psi \bar{e}_{\theta}, \quad \frac{\partial \bar{e}_{\psi}}{\partial \theta}=\cos \psi \bar{e}_{\theta}, \quad \frac{\partial \bar{e}_{\theta}}{\partial \theta}=-\sin \psi \overline{\boldsymbol{e}}_{r}-\cos \psi \overline{\boldsymbol{e}}_{\psi}
\end{gather*}
$$

Now that derivatives of the unit vectors have been characterized, we may proceed to carry out the derivatives in Eq. (A.1.11). Some terms are zero because the unit vectors form an orthonormal set, and others vanish because some partial derivatives of a unit vector are zero. What remains leads to

$$
\begin{equation*}
\nabla^{2} p=\frac{\partial^{2} p}{\partial r^{2}}+\frac{2}{r} \frac{\partial p}{\partial r}+\frac{1}{r^{2}} \frac{\partial^{2} p}{\partial \psi^{2}}+\frac{\cot \psi}{r^{2}} \frac{\partial p}{\partial \psi}+\frac{1}{r^{2}(\sin \psi)^{2}} \frac{\partial^{2} p}{\partial \theta^{2}} \tag{A.1.16}
\end{equation*}
$$

## A.1.4 Velocity and Acceleration

To describe the velocity and acceleration of a particle, it is necessary to account for the time dependence of the spherical coordinates. The position of a particle depends explicitly on the value of $r$ because $\bar{x}=r \bar{e}_{r}$. However, the position also depends on $\psi$ and $\theta$ because the $\bar{e}_{r}$ depends on $\psi$ and $\theta$. Accordingly, differentiating position requires application of the chain rule,

$$
\begin{equation*}
\bar{v} \equiv \frac{d \bar{x}}{d t}=\dot{r} \bar{e}_{r}+r \frac{\partial \bar{e}_{r}}{\partial r} \dot{r}+r \frac{\partial \bar{e}_{r}}{\partial \psi} \dot{\psi}+r \frac{\partial \bar{e}_{r}}{\partial \theta} \dot{\theta} \tag{A.1.17}
\end{equation*}
$$

Equation (A.1.15) describe the derivatives of the unit vectors, so the velocity is

$$
\begin{equation*}
\bar{v}=\dot{r} \bar{e}_{r}+r \dot{\psi} \bar{e}_{\psi}+r \dot{\theta} \sin \psi \bar{e}_{\theta} \tag{A.1.18}
\end{equation*}
$$

An expression for the acceleration is derived similarly. However, the motion variable of interest in fluid motion is the particle velocity, so we consider the velocity components to be functions of time that are known. For this reason, the preceding is rewritten as

$$
\begin{gather*}
\bar{v}=v_{r} \bar{e}+v_{\psi} \bar{e}_{\psi}+v_{\theta} \bar{e}_{\theta} \\
\dot{r}=v_{r}, \quad \dot{\psi}=\frac{v_{\psi}}{r}, \quad \dot{\theta}=\frac{v_{\theta}}{r \sin \psi} \tag{A.1.19}
\end{gather*}
$$

Differentiation of this description of velocity leads to

$$
\begin{align*}
\bar{a}= & \dot{v}_{r} \bar{e}_{r}+\dot{v}_{\psi} \bar{e}_{\psi}+\dot{v}_{z} \bar{e}_{z}+v_{r}\left(\frac{\partial \bar{e}_{r}}{\partial r} \dot{r}+\frac{\partial \bar{e}_{r}}{\partial \psi} \dot{\psi}+\frac{\partial \bar{e}_{r}}{\partial \theta} \dot{\theta}\right) \\
& +v_{\psi}\left(\frac{\partial \bar{e}_{\psi}}{\partial r} \dot{r}+\frac{\partial \bar{e}_{\psi}}{\partial \psi} \dot{\psi}+\frac{\partial \bar{e}_{\psi}}{\partial \theta} \dot{\theta}\right)+v_{\theta}\left(\frac{\partial \bar{e}_{\theta}}{\partial r} \dot{r}+\frac{\partial \bar{e}_{\theta}}{\partial \psi} \dot{\psi}+\frac{\partial \bar{e}_{\theta}}{\partial \theta} \dot{\theta}\right) \tag{A.1.20}
\end{align*}
$$

The rates of change of the spherical coordinates are replaced by the velocity components in Eq. (A.1.19), and Eq.(A.1.15) are used to describe how the unit vectors change. Collection of like components yields

$$
\begin{align*}
\bar{a}= & \dot{v}_{r} \bar{e}_{r}+\dot{v}_{\psi} \bar{e}_{\psi}+\dot{v}_{\theta} \bar{e}_{\theta}+v_{r}\left(\frac{v_{\psi}}{r} \bar{e}_{\psi}+\frac{v_{\theta}}{r} \bar{e}_{\theta}\right) \\
& +v_{\psi}\left(-\frac{v_{\psi}}{r} \bar{e}_{r}+\frac{v_{\theta}}{r \sin \psi} \cos \psi \bar{e}_{\theta}\right)+v_{\theta} \frac{v_{\theta}}{r \sin \psi}\left(-\sin \psi \bar{e}_{r}-\cos \psi \bar{e}_{\psi}\right) \tag{A.1.21}
\end{align*}
$$

$$
\begin{equation*}
\bar{a}=\left(\dot{v}_{r}-\frac{v_{\psi}^{2}}{r}-\frac{v_{\theta}^{2}}{r}\right) \bar{e}_{r}+\left(\dot{v}_{\psi}-\frac{v_{r}^{2}}{r}-\frac{v_{\theta}^{2}}{r} \cot \psi\right) \bar{e}_{\psi}+\left(\dot{v}_{\theta}+\frac{v_{r} v_{\theta}}{r}\right) \bar{e}_{\theta} \tag{A.1.22}
\end{equation*}
$$

The primary use of Eq. (A.1.18) for acoustics is to formulate a velocity continuity condition at an interface, and to evaluate the intensity. There are few contexts in which the acceleration arises, so Eq. (A.1.22) seldom is needed.

## A. 2 Cylindrical Coordinates

Cylindrical coordinates are an extension of polar coordinates in which the distance from the plane, which is the axial distance $z$, provides the third coordinate. Thus, if a variable only occurs in the $x y$ plane, or if a three-dimensional field variable is constant in the axial direction, setting all derivatives with respect to $z$ will convert the formulas derived here to polar coordinates.

## A.2.1 Transformations

The definition of cylindrical coordinates begins with the selection of a reference direction to be the axis of the coordinate system. In most cases, we are free to label directions as we wish, so we shall label this reference axis as $z$. The plane for the polar coordinates is defined such that $z$ is perpendicular to it. The polar coordinates are $R$ and $\theta$, but here too, other labels might be used. In Fig. A.2, $\theta$ is defined relative to the $x$-axis in the sense of the right hand rule. (Point the thumb of the right hand parallel to the $z$-axis. The curled fingers of that hand define the sense in which $\theta$ increases.) This rule is somewhat arbitrary, for example, one might wish to define $\theta$ relative to the $y$-axis, but the definition in the figure will suffice for our purposes.

The coordinate transformations between $(x, y, z)$ Cartesian coordinates and $(R . \theta, z)$ cylindrical coordinates may be derived by projecting a point onto the coordinate planes. Figure A. 2 shows that the axial distance $z$ locates a point above or below the $x y$ plane; it is the same as the Cartesian coordinate. In a three-dimensional situation, we should think of $R$ as the perpendicular distance from the point to the $z$-axis. Thus, we say that it is the transverse distance. In addition to being descriptive, using the adjective transverse aids us to avoid confusing it with the radial distance $r$ in spherical coordinates, which is distance to a reference point. The angle $\theta$ measures how much the transverse line is rotated about the $z$-axis. It is the circumferential angle because it locates a point along the circumference of the circle whose radius is $R$.

Trigonometry shows that the forward transformation is

$$
\begin{gather*}
x=R \cos \theta \\
y=R \sin \theta  \tag{A.2.23}\\
z=z
\end{gather*}
$$

We may determine the inverse transformation either by solving the above, or else through another trigonometric analysis. The result is

Fig. A. 2 Definition of cylindrical coordinates when $z$ defines the axial direction


$$
\begin{gather*}
R=\left(x^{2}+y^{2}\right)^{1 / 2} \\
\theta=\cos ^{-1}\left(\frac{x}{\left(x^{2}+y^{2}\right)^{1 / 2}}\right)=\sin ^{-1}\left(\frac{y}{\left(x^{2}+y^{2}\right)^{1 / 2}}\right)  \tag{A.2.24}\\
z=z
\end{gather*}
$$

It should be noted that the circumferential angle has not been described as $\theta=$ $\tan ^{-1}(y / x)$ because doing so leads to ambiguity in the assignment of quadrant.

Every curvilinear coordinate system has an associated set of unit vectors. They describe the direction in which a point would displace if one coordinate were increased and the others held constant. The cylindrical coordinate directions are the transverse direction $\bar{e}_{R}$, the circumferential direction $\bar{e}_{\theta}$, and the axial direction $\bar{e}_{z}$. Increasing $R$ displaces the point outward perpendicularly to the $z$-axis, and increasing $z$ displaces the point parallel to the $z$-axis. When $\theta$ is increased with $R$ and $z$ held constant, the point moves along a circle of radius $R$ parallel to the $x y$ plane. Thus, $\bar{e}_{R}, \bar{e}_{\theta}$, and $\bar{e}_{z}$ form a mutually orthogonal set of unit vectors. A cylindrical coordinate description of a vector entails finding the projections of the vector along each of these unit vectors. In turn, these unit vectors may be defined as components relative to the Cartesian coordinate directions. Reference to Fig. A. 2 shows that

$$
\begin{gather*}
\bar{e}_{R}=\bar{e}_{x} \cos \theta+\bar{e}_{y} \sin \theta \\
\bar{e}_{\theta}=-\bar{e}_{x} \sin \theta+\bar{e}_{y} \cos \theta  \tag{A.2.25}\\
\bar{e}_{z}=\bar{e}_{z}
\end{gather*}
$$

The orthogonal nature of the cylindrical coordinate directions may be verified by using the preceding to evaluate $\bar{e}_{R} \cdot \bar{e}_{\theta}, \bar{e}_{R} \cdot \bar{e}_{z}$, and $\bar{e}_{\theta} \cdot \bar{e}_{z}$.

## A.2.2 Gradient

The derivation of $\nabla p$ in terms of cylindrical coordinates follows the general procedure for the spherical coordinate derivation. Because $z$ is the same for cylindrical and Cartesian coordinates and $z$ is independent of $x$ and $y$, some terms in the chain rule for differentiation with respect to $x, y$, and $z$ are identically zero. The nonzero terms that result when $p$ is a function of $R, \theta$, and $z$ are

$$
\begin{align*}
\frac{\partial p}{\partial x}= & \frac{\partial p}{\partial R} \frac{\partial R}{\partial x}+\frac{\partial p}{\partial \theta} \frac{\partial \theta}{\partial x} \\
\frac{\partial p}{\partial y}= & \frac{\partial p}{\partial R} \frac{\partial R}{\partial y}+\frac{\partial p}{\partial \theta} \frac{\partial \theta}{\partial y}  \tag{A.2.26}\\
& \frac{\partial p}{\partial z}=\frac{\partial p}{\partial z}
\end{align*}
$$

We could evaluate the derivatives of the spherical coordinates by differentiating Eq. (A.2.24). However, an indirect approach that differentiates Eq. (A.2.23) is more
efficient, because the resulting derivatives only depend on the spherical coordinate. The procedure forms

$$
\begin{align*}
& 1=\frac{\partial R}{\partial x} \cos \theta-R \frac{\partial \theta}{\partial x} \sin \theta, \quad 0=\frac{\partial R}{\partial y} \cos \theta-R \frac{\partial \theta}{\partial y} \sin \theta  \tag{A.2.27}\\
& 0=\frac{\partial R}{\partial x} \sin \theta+R \frac{\partial \theta}{\partial x} \cos \theta, \quad 1=\frac{\partial R}{\partial y} \sin \theta+R \frac{\partial \theta}{\partial y} \cos \theta
\end{align*}
$$

The matrix form of these equations is

$$
\left[\begin{array}{cc}
\cos \theta & -R \sin \theta  \tag{A.2.28}\\
\sin \theta & R \cos \theta
\end{array}\right]\left[\begin{array}{ll}
\frac{\partial R}{\partial x} & \frac{\partial R}{\partial y} \\
\frac{\partial \theta}{\partial x} & \frac{\partial \theta}{\partial y}
\end{array}\right]=\left[\begin{array}{ll}
1 & 0 \\
0 & 1
\end{array}\right]
$$

Inversion of the coefficient matrix gives

$$
\left[\begin{array}{ll}
\frac{\partial R}{\partial x} & \frac{\partial R}{\partial y}  \tag{A.2.29}\\
\frac{\partial \theta}{\partial x} & \frac{\partial \theta}{\partial y}
\end{array}\right]=\left[\begin{array}{cc}
\cos \theta & \sin \theta \\
-\frac{1}{R} \sin \theta & \frac{1}{R} \cos \theta
\end{array}\right]
$$

Substitution of these derivatives into Eq. (A.2.26), followed by assembly of the individual terms to form $\nabla p$ leads to

$$
\begin{align*}
\nabla p= & \left(\frac{\partial p}{\partial R} \cos \theta-\frac{\partial p}{\partial \theta} \frac{1}{R} \sin \theta\right) \bar{e}_{x} \\
& +\left(\frac{\partial p}{\partial R} \sin \theta+\frac{\partial p}{\partial \theta} \frac{1}{R} \cos \theta\right) \bar{e}_{y}  \tag{A.2.30}\\
& +\frac{\partial p}{\partial z} \bar{e}_{z}
\end{align*}
$$

When we collect the coefficients of each derivative of $p$, we obtain

$$
\begin{equation*}
\nabla p=\frac{\partial p}{\partial R}\left(\bar{e}_{x} \cos \theta+\bar{e}_{y} \sin \theta\right)+\frac{\partial p}{\partial \theta}\left(-\frac{\bar{e}_{x}}{R} \sin \theta+\frac{\bar{e}_{y}}{R} \cos \theta\right)+\frac{\partial p}{\partial z} \bar{e}_{z} \tag{A.2.31}
\end{equation*}
$$

A comparison of these coefficients to Eq. (A.2.25) reveals that

$$
\begin{equation*}
\nabla p=\frac{\partial p}{\partial R} \bar{e}_{R}+\frac{1}{R} \frac{\partial p}{\partial \theta} \bar{e}_{\theta}+\frac{\partial p}{\partial z} \bar{e}_{z} \tag{A.2.32}
\end{equation*}
$$

## A.2.3 Laplacian

The procedure by which we will describe $\nabla^{2} P \equiv \nabla \cdot \nabla P$ is to use Eq. (A.2.32).This will entail describing the derivatives of the cylindrical coordinate unit vectors. The axial direction $\bar{e}_{z}$ is constant, and Eq. (A.2.25) indicates that $\bar{e}_{R}$ and $\bar{e}_{\theta}$ depend solely on $\theta$. Therefore, the derivatives are

$$
\begin{gather*}
\frac{\partial \bar{e}_{R}}{\partial R}=0, \quad \frac{\partial \bar{e}_{R}}{\partial \theta}=-\bar{e}_{x} \sin \theta+\bar{e}_{y} \cos \theta=\bar{e}_{\theta}, \quad \frac{\partial \bar{e}_{R}}{\partial z}=0 \\
\frac{\partial \bar{e}_{\theta}}{\partial R}=0, \quad \frac{\partial \bar{e}_{\theta}}{\partial \theta}=\bar{e}_{x} \cos \theta+\bar{e}_{y} \sin \theta=-\bar{e}_{R}, \quad \frac{\partial \bar{e}_{\theta}}{\partial z}=0  \tag{A.2.33}\\
\frac{\partial \bar{e}_{z}}{\partial R}=\frac{\partial \bar{e}_{z}}{\partial \theta}=\frac{\partial \bar{e}_{z}}{\partial z}=0
\end{gather*}
$$

The next step is to use Eq. (A.2.32) to describe both the divergence operator and $\nabla P$. The operation requires that the derivatives be applied to each component of $\nabla p$ before the dot product is applied. Thus, the operations leading to the Laplacian are

$$
\begin{align*}
\nabla^{2} p \equiv \nabla \cdot \nabla p= & \bar{e}_{R} \cdot \frac{\partial}{\partial R}\left(\frac{\partial p}{\partial R} \bar{e}_{R}+\frac{1}{R} \frac{\partial p}{\partial \theta} \bar{e}_{\theta}+\frac{\partial p}{\partial z} \bar{e}_{z}\right) \\
& +\frac{1}{R} \bar{e}_{\theta} \cdot \frac{\partial}{\partial \theta}\left(\frac{\partial p}{\partial R} \bar{e}_{R}+\frac{1}{R} \frac{\partial p}{\partial \theta} \bar{e}_{\theta}+\frac{\partial p}{\partial z} \bar{e}_{z}\right)  \tag{A.2.34}\\
& +\bar{e}_{z} \cdot \frac{\partial}{\partial z}\left(\frac{\partial p}{\partial R} \bar{e}_{R}+\frac{1}{R} \frac{\partial p}{\partial \theta} \bar{e}_{\theta}+\frac{\partial p}{\partial z} \bar{e}_{z}\right)
\end{align*}
$$

The unit vectors are mutually orthogonal, and they only depend on $\theta$. Thus, the nonzero terms arising from the preceding are

$$
\begin{equation*}
\nabla^{2} p=\frac{\partial}{\partial R}\left(\frac{\partial p}{\partial R}\right) \bar{e}_{R} \cdot \bar{e}_{R}+\frac{1}{R} \frac{\partial p}{\partial R} \bar{e}_{\theta} \cdot \frac{\partial \bar{e}_{R}}{\partial \theta}+\frac{1}{R^{2}} \frac{\partial}{\partial \theta}\left(\frac{\partial p}{\partial \theta}\right) \bar{e}_{\theta} \cdot \bar{e}_{\theta}+\frac{\partial^{2} p}{\partial z^{2}} \bar{e}_{z} \cdot \bar{e}_{z} \tag{A.2.35}
\end{equation*}
$$

The result of substituting Eq. (A.2.33) is

$$
\begin{equation*}
\nabla^{2} p=\frac{\partial^{2} p}{\partial^{2} R}+\frac{1}{R} \frac{\partial p}{\partial R}+\frac{1}{R^{2}} \frac{\partial^{2} p}{\partial \theta^{2}}+\frac{\partial^{2} p}{\partial z^{2}} \tag{A.2.36}
\end{equation*}
$$

Some individuals prefer to write the Laplacian in a manner that groups the derivatives with respect to $R$,

$$
\begin{equation*}
\nabla^{2} p=\frac{1}{R} \frac{\partial}{\partial R}\left(R \frac{\partial p}{\partial R}\right)+\frac{1}{R^{2}} \frac{\partial^{2} p}{\partial \theta^{2}}+\frac{\partial^{2} p}{\partial z^{2}} \tag{A.2.37}
\end{equation*}
$$

Other than brevity, there is little advantage to this form. For example, there is no power $n$ for which $R^{n} f(t-R / c)$ satisfies the wave equation in a situation where $p$ is independent of $\theta$.

## A.2.4 Velocity and Acceleration

The derivation of velocity and acceleration in terms of cylindrical coordinates follows the analysis for spherical coordinate. The first step is to construct the position by examining Fig. A.2. The field point is reached from the origin by traveling $R$ units in the transverse direction and $z$ units in the axial direction. The angle $\theta$ implicitly affects this construction because it defines the orientation of $\bar{e}_{R}$. Thus

$$
\begin{equation*}
\bar{x}=R \bar{e}_{R}+z \bar{e}_{z} \tag{A.2.38}
\end{equation*}
$$

The axial unit vector points in a constant direction and $\bar{e}_{R}$ depends on $\theta$, so differentiation of the position gives

$$
\begin{equation*}
\bar{v}=\dot{R} \bar{e}_{R}+\dot{z} \bar{e}_{z}+R \frac{\partial \bar{e}_{R}}{\partial \theta} \dot{\theta} \tag{A.2.39}
\end{equation*}
$$

The derivative of $\bar{e}_{R}$ is described in Eq. (A.2.33). The result is that the velocity is

$$
\begin{equation*}
\bar{v}=\dot{R} \bar{e}_{R}+R \dot{\theta}_{\theta}+\dot{z} \bar{e}_{z} \tag{A.2.40}
\end{equation*}
$$

In terms of the velocity components, the preceding is

$$
\begin{gather*}
\bar{v}=v_{R} \bar{e}_{R}+v_{\theta} \bar{e}_{\theta}+v_{z} \bar{e}_{z} \\
v_{R}=\dot{R}, \quad v_{\theta}=R \dot{\theta}, \quad v_{z}=\dot{z} \tag{A.2.41}
\end{gather*}
$$

To describe the acceleration, only the velocity components are considered to depend explicitly on $t$, whereas the unit vectors depend on time implicitly through their dependence on $\theta$. Hence, differentiation of the velocity expression leads to

$$
\begin{equation*}
\bar{a}=\dot{v}_{R} \bar{e}_{R}+\dot{v}_{\theta} \bar{e}_{\theta}+\dot{v}_{z} \bar{e}_{z}+v_{R} \frac{\partial \bar{e}_{R}}{\partial \theta} \dot{\theta}+v_{\theta} \frac{\partial \bar{e}_{\theta}}{\partial \theta} \dot{\theta} \tag{A.2.42}
\end{equation*}
$$

Setting $\dot{\theta}=v_{\theta} / R$ and collecting like components lead to the acceleration being described by

$$
\begin{equation*}
\bar{a}=\left(\dot{v}_{R}-\frac{v_{\theta}^{2}}{R}\right) \bar{e}_{R}+\left(\dot{v}_{\theta}+\frac{v_{R} v_{\theta}}{R}\right) \bar{e}_{\theta}+\dot{v}_{z} \bar{e}_{z} \tag{A.2.43}
\end{equation*}
$$

## Appendix B Fourier Transforms

A Fourier transform may be used to represent a function whose independent variable extends from minus to plus infinity. A corollary is that a Fourier transform is suitable for acoustic models of an infinite plane or an infinitely long cylinder. In both cases, the Fourier transform may be used to describe the dependence of the field on coordinates that measure position parallel to the boundary. In the time domain, the range for use of a Fourier transform is $-\infty<t<\infty$. However, we usually take $t=0$ to be the initial time. We may bypass the range requirement in this case by defining the function to be zero for $t<0$, but doing so requires that we not consider initial conditions. Moreover, response in the time domain is readily handled by standard methods, as well the Laplace transforms, which is derived from the Fourier transform. For this reason, the focus here is application of Fourier transforms to functions of position.

## B. 1 Derivation

We begin with a function $p(z)$ that is periodic in a distance $\ell$. The Fourier transform will emerge as $\ell \rightarrow \infty$. The complex amplitude of a plane wave propagating in the direction of increasing $z$ has been represented as $e^{-i k z}$, so we use the same convention for the complex exponential of the Fourier series,

$$
\begin{equation*}
p(z)=\frac{1}{2} \sum_{m=-\infty}^{\infty} P_{m} e^{-i m 2 \pi z / \ell} \tag{B.1.1}
\end{equation*}
$$

Because we are dealing with a spatial function, we say that $\ell$ is the wavelength and $2 \pi / \ell$ is the fundamental wavenumber. The Fourier coefficients corresponding to a specified function $p(z)$ are found by evaluating

$$
\begin{equation*}
P_{m}=\frac{2}{\ell} \int_{-\ell / 2}^{\ell / 2} p(z) e^{i m 2 \pi z / \ell} d z \tag{B.1.2}
\end{equation*}
$$

To obtain the limiting behavior as $\ell \rightarrow \infty$ we denote the wavenumber of the $m$ th harmonic as $\mu_{m}=m 2 \pi / \ell$ and note that the separation between adjacent wavenumbers is $\Delta \mu=2 \pi / \ell$. In terms of these parameters, the coefficients are given by

$$
\begin{equation*}
P_{m}=\frac{\Delta \mu}{\pi} \int_{-\ell / 2}^{\ell / 2} p(z) e^{i \mu_{m} z} d z \tag{B.1.3}
\end{equation*}
$$

For a given $p(z)$ and $\ell$, the quantity obtained from the integral is solely a function of $\mu_{m}$. Let us designate this integral as $\tilde{P}\left(\mu_{m}\right)$. Accordingly, we shall replace $P_{m}$ with $(\Delta \mu / \pi) \tilde{P}\left(\mu_{m}\right)$. This changes the Fourier series to

$$
\begin{equation*}
p(z)=\frac{1}{2 \pi} \sum_{m=-\infty}^{\infty} \tilde{P}\left(\mu_{m}\right) e^{-i m 2 \pi z / \ell} \Delta \mu \tag{B.1.4}
\end{equation*}
$$

Increasing $\ell$ decreases $\Delta \mu$, but $\Delta \mu$ is the wavenumber spacing between adjacent values of $\tilde{P}\left(\mu_{m}\right)$. Hence, a plot of $\tilde{P}\left(\mu_{m}\right)$ against its associated value of $\mu_{m}$ approaches a curve describing $\tilde{P}(\mu)$ as a function of $\mu$. This function is the Fourier transform of $p(z)$. A variety of notations have been used to denote the operation of Fourier transforming a function. Our choice is $\mathcal{F}(p(z), \mu)$, which explicitly lists the function that is transformed and the transform variable. Hence, we find that

$$
\begin{equation*}
\mathcal{F}(p(z), \mu) \equiv \tilde{P}(\mu)=\int_{-\infty}^{\infty} p(z) e^{i \mu z} d z \tag{B.1.5}
\end{equation*}
$$

To determine $p(z)$ when $\tilde{P}(\mu)$ is known we return to Eq. (B.1.4). Because $\Delta \mu$ becomes $d \mu$ as $\ell \rightarrow \infty$, the summation becomes an integral. The result is the inverse Fourier transform. This operation is designated as $\mathcal{F}^{-1}(\tilde{P}(\mu), z)$, so we have

$$
\begin{equation*}
\mathcal{F}^{-1}(\tilde{P}(\mu), z) \equiv p(z)=\frac{1}{2 \pi} \int_{-\infty}^{\infty} \tilde{P}(\mu) e^{-i \mu z} d \mu \tag{B.1.6}
\end{equation*}
$$

This is a heuristic derivation, but there is a formal mathematical foundation. ${ }^{1}$ Not all functions have a Fourier transform. A sufficient condition for its evaluation is that the integral of $|p(z)|$ be finite. No periodic function, such as a sine or cosine, meets this requirement, but we will see in the next section that it is possible to use Fourier transforms to deal with such functions. The inverse Fourier transform requires values of $\tilde{\mathcal{P}}(\mu)$ covering the full wavenumber spectrum $-\infty<\mu<\infty$. However, if the original function $p(z)$ is real, then $\tilde{P}(-\mu)=\tilde{P}(\mu)^{*}$, so only the positive spectrum needs to be evaluated.

[^82]At the outset, we restricted our attention to the Fourier transformation of a spatial function based on the usage of $\exp (-i \mu z)$ for a wave propagating in the positive $z$ direction. However, it is quite likely that you will eventually encounter a work in which the analysis is based on using $\exp (i \mu z)$. Furthermore, it might be appropriate to employ the Fourier transform to a describe function of time based, for which we have adopted $\exp (i \omega t)$ for harmonic functions. Both situations require a modification of the definition of a Fourier transform and its inverse. Let us denote this alternative definition as $\Phi(p(\xi), \mu)$ where $\xi$ is the independent space or time variable and $\mu$ is the transform parameter. This alternative is

$$
\begin{align*}
\Phi(p(\xi), \mu) & \equiv \hat{P}(\mu)=\int_{-\infty}^{\infty} p(\xi) e^{-i \mu \xi} d \xi \\
\Phi^{-1}(\hat{P}(\mu), \xi) & \equiv p(\xi)=\frac{1}{2 \pi} \int_{-\infty}^{\infty} \hat{P}(\mu) e^{i \mu \xi} d \mu \tag{B.1.7}
\end{align*}
$$

When it is necessary to distinguish between the alternatives, we will refer to Eq.(B.1.5) as the spatial Fourier transform, and Eq.(B.1.7) as the temporal Fourier transform.

There is much similarity of this alternative definition to the one with which we began. Therefore, it should not be surprising that given one, the other is readily recovered. The following relations are readily verified

$$
\begin{gather*}
\Phi(p(\xi),-\mu)=\mathcal{F}(p(\xi), \mu)  \tag{B.1.8}\\
\Phi(p(\xi), \mu)^{*}=\mathcal{F}\left(p(\xi)^{*}, \mu\right), \quad \mathcal{F}(p(\xi), \mu)^{*}=\Phi\left(p(\xi)^{*}, \mu\right)
\end{gather*}
$$

Variants on the present Fourier transform definition multiply the integral by different coefficients. One that is often used leads to a transform and its inverse that have similar forms. This quality is obtained if the integral in Eq.(B.1.7) or (B.1.5) is multiplied by $1 /(2 \pi)^{1 / 2}$, which also becomes the factor multiplying the integral for the inverse transform. Another version, which is not often used, has a $1 /(2 \pi)$ factor for the transform and none for the inverse. Knowledge of the relationships between the various definitions is needed when we wish to use a work that has a definition that differs from our selection. Equation (B.1.8) can be quite useful in this regard because it allows us to employ any set of tables of functions and their Fourier transforms.

## B. 2 Evaluation Techniques

Our concern here is evaluation of the Fourier transform of a spatial function $p(z)$, as well as evaluation of the spatial function when we know a transform $\tilde{P}(\lambda)$. If $p(z)$ is not too complicated, we can evaluate the integrals directly, possibly with the aid of a table of integrals. However, there are alternative methods of evaluation that are
likely to require less effort than that entailed for a formal integration. One is based on using tabulations of transform pairs rather than of integrals. The other makes use of FFTs.

## B.2.1 Transform Pairs

If a function $p(z)$ is so familiar that its Fourier integral can be obtained with the aid of a table of integrals, it is likely that the result appears in a table of Fourier transform pairs. An entry in such a tabulation An entry in such a tabulation will list a function $p(z)$ and the adjacent entry will list the corresponding Fourier transform $\tilde{P}(\mu)$. Such a tabulation may be used in either direction: from the physical variable that is a function of $z$ to the transformed function of $\mu$, or vice versa. (Here we are using $z$ as the independent variable and $\mu$ as the transform parameter. Obviously, other symbols may be used interchangeably.) The first consideration in using a table of Fourier transforms is to ascertain that it is based on a definition of a Fourier transform that is consistent with one's work. If not, then the tabulation will require conversion as described in the previous section.

Extensive tables of Fourier transforms are available in several books, ${ }^{2}$ (1954), Ch. III. as well as some web sites. A condensed table, whose entries tend to be relevant to acoustic systems, are listed here. Table B. 1 gives some basic properties that tell us how to adapt a known transform pair $p(z)$ and $\tilde{P}(\mu)$ to handle related functions. The last entry in Table B. 1 is known as the convolution integral. It seldom is used to find a transform, but it is quite relevant for performing the inverse. One reason for this is that it fits a situation in which $\tilde{P}(\mu)$ is the transform of an excitation and $\tilde{Q}(\mu)$ is a transfer function. However, we will see in the next section that an FFT provides a much simpler alternative for evaluating a convolution integral.

Table B. 2 lists some specific functions of $z$ and their transform. Several entries contain the Dirac delta function $\delta(\mu)$. In each case, the associated function $p(z)$ does not meet the specification that the integral of $|p(z)|$ be finite, which is why the transform is singular. One way to verify these transforms is to evaluate the corresponding $p(z)$ function as an inverse transform. For example, to verify the entry for a complex exponential we observe that the integrand for the inverse transform is $2 \pi \delta\left(\mu-\mu_{0}\right) e^{-i \mu z}$ over an infinite range of $\mu$ yields $2 \pi e^{-i \mu z}$ evaluated at $\mu=\mu_{0}$, that is, $2 \pi e^{-i \mu_{0} z}$. The integral for the inverse transform is multiplied by $1 /(2 \pi)$, so we have shown that $\mathcal{F}^{-1}\left(2 \pi \delta\left(\mu-\mu_{0}\right), z\right)=\exp \left(-i \mu_{0} z\right)$.

The entries in Table B. 2 may be extended by using the properties in TableB.1. For example, suppose we wish to transform $\sin (k z)$. The Euler identity and linearity leads to

$$
\begin{equation*}
\mathcal{F}(\sin (k z), \mu) \equiv\left(\frac{1}{2 i}\right) \mathcal{F}\left(e^{i k z}-e^{-i k z}, \mu\right)=\frac{\pi}{i}[\delta(\mu+k)-\delta(\mu-k)] \tag{B.2.9}
\end{equation*}
$$

[^83]Table B. 1 Basic properties relating an arbitrary function and its Fourier transform

| Name | Function | Fourier transform |
| :--- | :--- | :--- |
| Definition | $p(z)=\frac{1}{2 \pi} \int_{-\infty}^{\infty} \tilde{P}(\mu) e^{-i \mu z} d \mu$ | $\tilde{P}(\mu)=\int_{-\infty}^{\infty} p(z) e^{i \mu z} d z$ |
| Linearity | $a p(z)+b q(z)$ | $a \tilde{P}(z)+b \tilde{Q}(z)$ |
| Coordinate scaling | $p(\alpha z)$ | $\frac{1}{\|\alpha\|} \tilde{P}\left(\frac{\mu}{\alpha}\right)$ |
| Coordinate shifting | $p\left(z-z_{0}\right)$ | $\tilde{P}(\mu) e^{i \mu z_{0}}$ |
| Modulation | $p(z) e^{-i \mu_{0} z}$ | $\tilde{P}\left(\mu-\mu_{0}\right)$ |
| Complex conjugate | $p(z)^{*}$ | $\tilde{P}(-\mu)^{*}$ |
| nth derivative | $\frac{d^{n}}{d z^{n}} p(z)$ | $(-i \mu)^{n} \tilde{P}(\mu)$ |
| Definite integration | $\int_{-\infty}^{z} p(\eta) d \eta$ | $-\frac{1}{i \mu} \tilde{P}(\mu)+\pi \tilde{P}(0) \delta(\mu)$ |
| Differentiate the <br> transform | $(i z)^{n} p(z)$ | $\frac{d^{n}}{d \mu^{n}} \tilde{P}(\mu)$ |
| Convolution | $\int_{-\infty}^{\infty} p(\xi) q(z-\xi) d \xi=$ | $\tilde{P}(\mu) \tilde{Q}(\mu)$ |

A particularly useful function is sometimes referred to as the "box" function. It is defined as $\operatorname{box}(z)=h(z-a)-h(z-b)$ with $b>a$. This function is one for $a<z<b$, zero outside that band, and one half at the discontinuities. We could find the corresponding transform by starting with the rect function, but application of the linearity and coordinate shifting properties to the definition is more direct. Thus,

$$
\begin{align*}
\mathcal{F}(\operatorname{box}(z), \mu) & =\mathcal{F}(h(z-a), \mu)-\mathcal{F}(h(z-b), \mu)=e^{i \mu a} \mathcal{F}(h(z), \mu)-e^{i \mu b} \mathcal{F}(h(z), \mu) \\
& =\left(e^{i \mu a}-e^{i \mu b}\right)\left[\pi \delta(\mu)-\frac{1}{i \mu}\right]=\frac{1}{i \mu}\left(e^{i \mu b}-e^{i \mu a}\right) \tag{B.2.10}
\end{align*}
$$

where the last form results from $\delta(\mu)$ being zero except at $\mu=0$.

Table B. 2 Common functions and their Fourier transform

| Descriptive name | Function | Fourier transform |
| :--- | :--- | :--- |
| Constant | 1 | $2 \pi \delta(\mu)$ |
| Step | $h(z)$ | $\pi \delta(\mu)-\frac{1}{i \mu}$ |
| Impulse | $\delta\left(z-z_{0}\right)$ | $e^{i \mu z_{0}}$ |
| Complex exponential | $e^{-i \mu 0 z}$ | $2 \pi \delta\left(\mu-\mu_{0}\right)$ |
| One-sided exponential decay | $e^{-\alpha z} h(z), \operatorname{Re}(\alpha)>0$ | $\frac{1}{(\alpha-i \mu)^{2}}$ |
| Attenuated ramp | $e^{-\alpha z} h(z)$ | $(2 \pi)^{1 / 2} z e^{-(\mu z)^{2} / 2}$ |
| Gaussian | $\operatorname{sinc}(\alpha z), \alpha>0$ | $\frac{1}{\alpha} \operatorname{rect}\left(\frac{\mu}{2 \pi \alpha}\right)$ |
| Normalized $\operatorname{sinc}:$ <br> sinc $(z) \equiv \frac{\sin (\pi z)}{\pi z}$ | $\operatorname{rect}\left(\frac{z}{z_{0}}\right)$ | $e^{-\alpha \operatorname{sinc}\left(\frac{\mu z_{0}}{2 \pi}\right)}$ |
| Rectangle: rect $(z) \equiv$ <br> $h\left(z+\frac{1}{2}\right)-h\left(z-\frac{1}{2}\right)$ | $\operatorname{Re}(\alpha)>0$ | $\frac{2 \alpha}{\alpha^{2}+\mu^{2}}$ |
| Symmetric exponential |  |  |

## B.2.2 Fast Fourier Transforms

For nonstandard functions analytical evaluation of a Fourier transform or inversion of a transform will require significant analytical effort, possibly involving contour integration in the complex space. If numerical results are sufficient, the Fast Fourier Transform (FFT) technology is ideal for that purpose. The presumption here is that the usage of Fast Fourier Transforms in conjunction with periodic time data, which was addressed in Sect. 1.4, is a familiar concept. The first objective here is to ascertain how an FFT computer routine may be used to evaluate the Fourier transform of an excitation whose spatial dependence is arbitrary. Then, we will examine how FFTs may be used in conjunction with transfer functions to evaluate an acoustical response.

Whether the Fourier transform of an arbitrary function $p(z)$ exists is a complicated issue. It is assured whether the function is zero outside some region centered around the origin.This is the situation for our studies, because the objects we are interested
in have finite extent. More generally, the transform will be finite if the function is absolutely integrable, that is,

$$
\begin{equation*}
\int_{-\infty}^{\infty}|f(x)| d x<\infty \tag{B.2.11}
\end{equation*}
$$

If we avoid processing periodic functions, this property will be satisfied for finite acoustical systems.

Usage of an FFT to treat nonperiodic data requires that we define a window whose width $L$ serves as the effective period. There are two considerations to selecting this window. The selection begins by identifying the interval $-\ell / 2<x<\ell / 2$ outside of which the data is negligible. For example, if the data is a Gaussian function, $p(z)=$ $a \exp \left(-\beta x^{2}\right)$, then the criterion $|p(z)|<\max (|p|) / 100$ if $|x|>\ell / 2$ would be met by setting $\ell / 2 \geq(\ln (100) / \beta)^{1 / 2}$. Data that suddenly drops off requires special care in this regard because the interval should be sufficient to capture the nature of such data. For example, if $p(z)$ is a rectangle function, then $\ell$ should be substantially larger than the width of the rectangle.

The value of $\ell$ that is identified should not be the period $L$. This prohibition stems from our ultimate objective of using transfer functions to determine the response to an excitation that has been processed with an FFT. This operation introduces wraparound error that narrows the interval in which the data is valid. It is possible to avoid contaminating the response with wraparound error. To do so, the length of the window must be at least twice $\ell$. The data is set to zero in the extended range. This is a process of zero padding. In other words, if the data is such that $|p(z)|<\varepsilon \max (|p|) / 100$ if $|x|>\ell / 2$, then the data window for the FFT would be $-L / 2 \leq x \leq L / 2$, with $L \geq 2 \ell$. In the intervals $-L / 2 \leq x<-\ell / 2$ and $\ell / 2<x \leq$ $L / 2$, we would set $p(z)=0$. If we follow this protocol, then FFT results will be valid for the original data interval, $-\ell / 2<x<\ell / 2$.

Let us first consider the formal definition of a DFT. After we have done so, we will examine how the definition is adjusted to a form that is suitable for an FFT algorithm. The period is $L$, and the data set consists of $N$ samples. Therefore, the discrete wavenumbers for the Fourier series are

$$
\begin{equation*}
\mu_{n}=n\left(\frac{2 \pi}{L}\right), n=-N / 2+1, \ldots, N / 2 \tag{B.2.12}
\end{equation*}
$$

Suppose we knew the function $p(z)$ underlying the sampled data. The Fourier transform of that function would integrate over $-L / 2<z<L / 2$, because we are considering $p(z)$ to be zero outside that interval. Thus, the Fourier transform at the series wavenumbers is

$$
\begin{equation*}
\mathcal{F}\left(p(z), \mu_{n}\right) \equiv \tilde{P}\left(\mu_{n}\right)=\int_{-L / 2}^{L / 2} p(z) e^{i \mu_{n} z} d z \equiv \int_{-L / 2}^{L / 2} p(z) e^{i(2 n \pi / L) z} d z \tag{B.2.13}
\end{equation*}
$$

Rather than attempting to carry out this integration analytically, let us consider a numerical approximation using a strip rule whose width is $\Delta z=L / N$. There are $N$ strips and the sample locations are $z_{j}=j \Delta z, j=-N / 2+1, \ldots, N / 2$. The resulting approximation is

$$
\begin{equation*}
\mathcal{F}\left(p(z), \mu_{n}\right) \approx \frac{L}{N} \sum_{j=-N / 2+1}^{N / 2} p_{j} e^{i 2 \pi j n / N}, n=-\frac{N}{2}+1, \ldots, \frac{N}{2} \tag{B.2.14}
\end{equation*}
$$

The significant aspect is that although the definition of a DFT is based on this approximation, the operation itself exists independently of the concept of a Fourier transform. Rather, it is a formal prescription of a way in which data may be processed. This distinction is indicated by using a slightly different notation. A "caret" will denote DFT data, whereas a "tilda" has been used for a Fourier transform. The definition of the DFT is

$$
\begin{equation*}
\hat{P}_{n} \equiv \sum_{j=-N / 2+1}^{N / 2} p_{j} e^{i 2 \pi j n / N}, n=-\frac{N}{2}+1, \ldots, \frac{N}{2} \tag{B.2.15}
\end{equation*}
$$

The inverse spatial DFT evaluates the $p\left(z_{j}\right)$ data given a set of DFT values. It is

$$
\begin{equation*}
p_{j}=\frac{1}{N} \sum_{n=-N / 2+1}^{N / 2} \hat{P}_{n} e^{-i 2 \pi n j / N}, \quad j=-\frac{N}{2}+1, \ldots, \frac{N}{2} \tag{B.2.16}
\end{equation*}
$$

One manifestation of the fact that a DFT is an independent operation, separate from a Fourier transform, is that the inverse operation is exact. That is, if we process a set of $p\left(z_{j}\right)$ data to obtain a DFT set $\hat{P}_{n}$, then take the inverse DFT of this data, the result will exactly be the $p\left(z_{j}\right)$ values (aside from numerical errors). No approximation is involved.

In order to implement an FFT algorithm and its inverse, the spatial data set and the DFT values are manipulated slightly. One alteration is that the spatial data is considered to be situated in the window $0 \leq z \leq L$, which constitutes a forward shift by half the fundamental wavelength. Shifting the spatial data leads to an algorithm that evaluates

$$
\begin{equation*}
P_{n}^{\prime}=\sum_{j=0}^{N-1} p_{j} e^{i 2 \pi j n / N}, n=0,1, \ldots, N-1 \tag{B.2.17}
\end{equation*}
$$

The notation $P_{n}^{\prime}$ is intended to distinguish the FFT output from the DFT values, which are denoted as $\hat{P}_{n}$. The shift of the data window is unimportant for our current purpose. In the event that it is necessary to compare the FFT output to the DFT values
obtained from the formal definition in Eq. (B.2.15), the coordinate shifting property in Table B. 1 for a shift of $L / 2$ multiplies the transform by $\exp \left(i \mu_{n} L / 2\right)=(-1)^{n}$.

The other adjustment arises because the range of $n$ appearing in the FFT definition above differs from the range of $n$ in Eq. (B.2.15). The FFT exploits the periodicity properties of a DFT to place the data for negative wavenumbers after the data for zero and positive wavenumbers. This is the complex conjugate mirror image property first described in Eq. (1.4.40). The rearrangement is described by

$$
\begin{gather*}
P_{n}^{\prime}=\hat{P}_{n}, \quad n=0, \ldots, N / 2  \tag{B.2.18}\\
P_{n}^{\prime}=\hat{P}_{n-N}, \quad n=N / 2+1, \ldots, N-1
\end{gather*}
$$

The rearrangement of wavenumbers follows the same rule

$$
\begin{gather*}
\mu_{n}^{\prime}=\mu_{n}=\frac{2 \pi n}{L}, \quad n=0, \ldots, N / 2  \tag{B.2.19}\\
\mu_{n}^{\prime}=\mu_{n-N}=\frac{2 \pi(n-N)}{L},
\end{gather*}
$$

This shift in the wavenumber spectrum is a crucial issue for the objective of using transfer functions to find a response.

Any routine that performs a Fast Fourier Transform will be accompanied by one that performs the inverse transform.FFT definition in Eq.(B.2.15) is

$$
\begin{equation*}
p_{n}=\frac{1}{N} \sum_{j=0}^{N-1} \hat{P}_{j} e^{-i 2 \pi n j / N}, \quad n=0,1, \ldots N-1 \tag{B.2.20}
\end{equation*}
$$

We will refer to a computer routine that implements the definition in Eq. (B.2.15) as a spatial FFT, whereas the FFT definition in Sect. 1.4 .4 will be said to be a a temporal FFT. (The only difference is the sign of the argument of the complex exponential.) If the available software package only has a routine that implements a temporal FFT, it will be necessary to adjust its output. One way to do so is to use the first of Eq. (B.1.8), which calls for swapping the negative and positive index of the DFT data. This can be somewhat problematic because of the arrangement of data for FFT algorithms, as described by Eq. (B.2.18). An alternative is to follow the last of Eq. (B.1.8) which calls for using the available FFT routine to transform the complex conjugate of the spatial data, then taking the complex conjugate of that output.

The preceding addresses the fundamentals of FFT processing. Let us now turn our attention to the reason that an FFT might be useful for acoustical system analysis. Suppose we have used an FFT algorithm to transform the spatial dependence of an excitation. Examples of this would be transformation of an arbitrary vibrational pattern along one direction of an infinite plane, or transformation of the axial pattern
of vibration on the surface of a circular cylinder. The transfer function in the first case would be the acoustic wave radiated into the fluid by a sinusoidal surface wave, which was determined in Chap. 5. In the case of a cylinder, an analysis in Chap. 7 found the wave radiated into the fluid by a helical wave surface wave at a fixed circumferential harmonic number. For both systems, the transfer function is the amplitude of a wave that is radiated due to an excitation that propagates along the surface in a certain direction, which we will designate as $x$. The wavenumber in that direction is designated as $\mu$. Let us denote the amplitude of the wave that propagates in the $x$ direction as $G(\mu)$. To be a transfer function, the excitation causing it must be a wave in the same direction with unit amplitude. Our task is to determine the signal when the excitation is an arbitrary function.

Let us analyze radiation from a plane to illustrate the process. Suppose the surface vibration is harmonic at frequency $\omega$. The distribution in the $x$ direction is arbitrary, such that the surface velocity normal to the plane is $\operatorname{Re}(v(x) \exp (i \omega t))$, where $v(x)$ may be complex. It is necessary that there is a DFT window $L$ that accommodates the significant portion of $v(x)$. The transfer function is the $y$ dependence of the wave radiated from the surface when $v(x)=\exp (-i \mu x)$. This quantity is described by Eq. (5.1.16), with the velocity amplitude $i \omega W$ set to one. Thus, the transfer function is such that

$$
\begin{gather*}
P=G(\mu) e^{-i \mu x}, \quad G(\mu)=\rho_{0} c \frac{k}{\kappa} e^{-i \kappa y} \\
\kappa=\left\{\begin{array}{l}
\left(k^{2}-\mu^{2}\right)^{1 / 2} \text { if } k>\mu \\
-i\left(\mu^{2}-k^{2}\right)^{1 / 2} \text { if } k<\mu
\end{array}\right. \tag{B.2.21}
\end{gather*}
$$

If the amplitude of the surface vibration is $V(\mu) \neq 1$, then the wave radiated from the surface would be $V(\mu) G(\mu) \exp (-i \mu x)$. Addition of the contribution of the full spectrum of $V(\mu)$ leads to

$$
\begin{equation*}
p(x, y, t)=\operatorname{Re}\left(P(x, y) e^{i \omega t}\right) \tag{B.2.22}
\end{equation*}
$$

where the complex pressure amplitude is given by

$$
\begin{equation*}
P(x, y)=\frac{1}{2 \pi} \int_{-\infty}^{\infty} V(\mu) G(\mu) e^{-i \mu x} d \mu=\int_{-\infty}^{\infty}\left(\rho_{0} c \frac{k(\mu)}{\kappa(\mu)} V(\mu) e^{-i \kappa(\mu) y}\right) e^{-i \mu x} d \mu \tag{B.2.23}
\end{equation*}
$$

This expression is an inverse Fourier transform. It is readily adapted to fit the definition of an inverse FFT in Eq. (B.2.20). First we evaluate the FFT of the excitation distribution $V(\mu)$, which produces velocity coefficients $V_{n}^{\prime}$. The FFT of the pressure is analogous to the integrand, specifically,

$$
\begin{equation*}
P_{n}^{\prime}=\rho_{0} c \frac{k}{\kappa_{n}} V_{n}^{\prime} e^{-i \kappa_{n} y}, \quad \kappa_{n}=\kappa\left(\mu_{n}\right) \tag{B.2.24}
\end{equation*}
$$

The last step is to take the inverse FFT of the $P_{n}^{\prime}$ data. The output of the inverse FFT will be values of $P(x, y)$ at fixed $y$ at the points $x_{n}$ used to sample $v(x)$. The only precaution for implementing this procedure is associated with the manner in which the DFT data is arranged in an FFT algorithm. Specifically, the first $N+1$ values are the associated with $\lambda_{n} \geq 0$. These values are followed by the data for $\lambda_{n}<0$. The wavenumbers corresponding to the data that is input to the inverse FFT are

$$
\mu_{n}=\left\{\begin{array}{l}
\frac{2 n \pi}{L}, \quad n=0,1, \ldots, \frac{N}{2}  \tag{B.2.25}\\
\frac{2(n-N) \pi}{L}, \quad n=\frac{N}{2}+1, \ldots, N-1
\end{array}\right.
$$

If the wavenumbers are arranged in this sequence as a vector array $\{\mu\}$ in a computer program, all operations may carried out vectorially, for example, kappa=conj (sqrt ( $\left.\mathrm{k}^{\wedge} 2-\mathrm{mu} .^{\wedge} 2\right)$ ) in MATLAB.

Although this FFT procedure is fairly straightforward, there is a profound difficulty that can inhibit its application for acoustics. The system we have set up consists of an interval along the boundary within which there is a vibration. This is the twodimensional analog of a piston in an infinite baffle. In both cases, increasing distance from the boundary leads to ever increasing spreading of the acoustic field parallel to the surface. The inverse FFT that generates this distribution must have a window width that accommodates the spreading effect. In other words, the window length must be selected on the basis of the extent of the signal at the selected distance $y$ to the boundary, rather than the nature of the surface vibration. This is where the difficulty arises, because the extent of the field is not known until the system has been analyzed.

## Index

## A

Absorbing boundary conditon, 123
Absorption coefficient, 607
Absorptive liner, 224
Acceleration
in cylindrical coordinates, 675
in spherical coordinates, 670
Acoustical imaging, 656
Acoustically compact, 99
Acoustic cavity, 85
Acoustic levitation, 655
Acoustic Mach number, 546, 553, 611
Acoustic perturbation, 409
Acoustic-structure interaction, 238
Added mass, 103, 501, see also Virtual mass
Adiabatic assumption, 541
Algorithm
for a modal series, 252, 264, 282, 308
for a nonlinear wave, 559, 563, 633
Ambient conditions, 409, 412
Amplitude
dispersion, 545, 645
slowly varying, 407, 428, 443
Angular spectrum, 176
Annular waveguide, 275
Antinode, 154
Anti-piston transducer, 156
Area function
for a multivalued waveform, 585
for the Webster horn equation, 191
Arrête, 417, 434
Asymptotic approximation
of a spherical harmonic series, 21
of Bessel functions, 58, 270, see also Farfield approximation
Atmosphere, propagation in, 406
Attenuation, 230, 521, 596, 607

Axial wavenumber, 57, 217, 267
complex, 229
Axisymmetric body, scattering by, 497
Axisymmetric field, 19, 61, 71
Azimuthal angle, 94, 140, 353
Azimuthal harmonic, 16, 268, 351

## B

Backscatter, 486, 502, 509, 511, 514
Backscatter cross section, 510, 512, 514, 517
Backward propagating wave, 199, 292, 298, 548
B/A coefficient, 551
Baffle, 29, 133, 160
Beam aspect, 97
Beamwidth, 148, 162
Beating signal
for group velocity, 202
from a supersonic source, 242
nonlinear, 607
Bending of rays, 420
Bessel function
cylindrical, 12, 57, 145
derivative of, 14, 270
for modes, 268, 271
integral for, 570
modified, 64, 273
of negative order, 62
spherical, 11, 15, 57, 312, 319
Bessel's equation, 57
modified, 64, 273
Biomedical applications, 657
Blocked area, 491, 513, 520
Blocked pressure, 522, 527
Blue sky, 491
Body force, 409

Born approximation, 485, 490
Boundary condition
for a cavity, 243, 300, 313
for a perturbation series, 614,639
for a waveguide, 188, 196, 214, 218
for farfield radiation, 125
mixed, 244
nonlinear, 638
Boundary elements, 111, 363
Box function, 681
Branch cut, 218, 226, 246
Breathing mode, 20
Bulk modulus, 40, 485, 491, 551
Burton-Miller formulation, 115

## C

Calculus of variations, 365, 463
Cauchy principal part, 109, 113
Causality, principle of, 451
Caustic, 417, 432
effect on amplitude, 450
example, 454
formed by nonlinear rays, 651
Cavitation, 577, 656
Cavity, 91, 93, 113, 291
Cavity mode, 293, 295, 297, 336, 365
Center of curvature, of a ray path, 423
Center wave, 153, 179
Chaladni lines, 346
Characteristic curve, 542, 544
Characteristic equation, 219, 225
for a cylinder, 325
for a spherical shell, 41
for a waveguide, 219, 224, 238, 259
CHIEF computer code, 114
Circular disk, scattered field, 508, 513
Circular ray path, 424, 443
Circumferential angle, 671
Circumferential harmonic, 69, 77, 271, 323
Coefficient of nonlinearity, 552
Coincidence frequency, 236
Compact set of sources, 18
Complex eigenvalues, 225
Complex wavenumber, 226
Compliant boundary, 273, 298, 326, 379, 386
Concave reflector, 512
Concentric cylinders, 352
Concentric spheres, 319,360
Conical horn, 188, 191
Connectivity matrix, 117

Conservation
of energy, 572, 579
of mass, 189, 541, 579
Constitutive nonlinearity, see Equation of state
Continuity condition
at a cavity wall, 361, 380
at a vibrating plate, 638
on the surface of an elastic shell, 39
Continuity equation, 409, 540, 598
Control volume, 190, 578
Convective nonlinearity, 544
Convergence
of a Legendre series, 8, 26
of a modal series, $42,47,250,262$
of a spherical harmonic series, 22, 526
of Rayleigh-Ritz method, 374
Coordinate straining, 621, 629, 642
Corner, for Kirchhoff-Helmholtz integral theorem, 110
Coupled equations
for a spherical shell, 40
for cavity-structure interaction, 385
for waveguide-structure interaction, 238
Creeping waves, 520
Critical incidence, 471
Cross-sectional area
of a horn, 189
of a ray tube, 445
Cumulative growth, 628
Curvature of rays, 423
Cutoff, 259
for an exponential horn, 194, 204
for radiation from a cylinder, 68
of waveguide modes, 218, 260
Cylindrical Bessel function, see Bessel function
Cylindrical coordinates, 56, 144, 267
definition of, 671
unit vectors, 672, 674
Cylindrical spreading, 63

## D

D'Alembert solution, 544, 548, 614
Decay constant, 220, 246
Decomposition of a propagating mode, 222
Degree of a Legendre function, 4, 16
Derivative of a unit vector
for cylindrical coordinates, 675
for spherical coordinates, 668
Differential scattering cross-section, 509, 517

Diffraction, 152, 473, 520
Dipole, 495
for scattering, 491
two-dimensional, 78
Dipole moment, 78, 100, 101
Dirac delta function, 295
Directivity
for a transducer in a baffle, 137, 144, 166
for radiation, 53, 97, 147
for scattering, 484, 517
from a spherical harmonic series, 21
of a horn, 187, 196
of a ring, 169
Disk
radiation from, 103
scattering from, 497, 501
Dispersion, 195, 202
Displacement
of a plate, 637
of a spherical shell, 37,522
structural, 383
Dissipation, 229, 567, 607
Divergence theorem, 86, 338, 340
Dome tweeter, 29
Dowell's approximation, 380

## E

Earnshaw solution, 540, 546
Edge wave, 154, 179
Eigenray, 458
Eigenvalue
complex, 225
for a rigid-walled cylinder, 218, 271
for a waveguide, 218, 225
purely imaginary, 228, 274
Eigenvalue problem, 337
Eikonal equation, 439, 440
Elastic plate, 230, 637
Elastic spherical shell, 36, 523
Elastic structure, 380
Elliptic cavity, 374
End-fire array, 657
Energy absorption, 300
Energy and work, 579
Energy functionals, 366
Energy, in the old-age stage, 594
Entropy, 581
Equal area rule, 582, 592
Equal eigenvalues, 261
Equal natural frequencies, 339, 345, 354
Equation of state
for a heterogeneous fluid, 410
for a liquid, 550
for an ideal gas, 549
relation to sound speed, 593
Equations of motion
for a spherical shell, 36
for a structure, 383
Error residual, 106
Euler's equation, 190, 381, 541, 578
Euler-Lagrange equation, 468
Evanescent mode, 222, 248, 260, 262
Evanescent wave, 195, 202
Exponential horn, 194
group velocity for, 204
Extensional elastic wave, 525

## F

Farfield
for a transducer, $127,133,144,149,154$, 163
for radiation from a cylinder, 69
for radiation from a spherical shell, 41, 52
of a nonlinear spherical wave, 632
of a spherical harmonic series, 21
of Kirchhoff-Helmholtz integral theorem, 94
of the Rayleigh integral, 144
Fast Fourier Transform, 682
for radiation, 84, 139
Fermat's principle, 463
FFT, see Fast Fourier Transform
Finite amplitude wave, see Nonlinear wave
Finite elements, 115, 363
Fitting shocks, 582, 587
Flexible wall, 42, 224, 236, 379
Fluid impedance, 44, 525
Fluid-loaded resonance, 51, 526
Fluid-loading effect of, 43
Fluid-structure interaction, 36, 243, 379
Focus, 431, 435
Focusing factor, 431
Force
acting on a sphere, 38
exerted on a piston, 166
for low frequency radiation, 102
Forward and backward waves
in a cavity, 292, 298
in a horn, 194, 199
Forward characteristic, 544
Forward scattering, 509, 520
Fourier coefficient, 677
Fourier series, 677
for a cylindrical waveguide, 267
for a nonlinear wave, 569, 603
for radiation from a cylinder, 56, 69
Fourier transform, 82, 138, 177, 677
alternative definitions, 679
conditions for, 678
of a surface velocity, 82, 139, 176
table of, 679, 680
Fubini-Ghiron, 571

## G

Galerkin method, 116, 118
Gamma function, 169
General eigenvalue problem, 373
Generalized coordinates, 383
Generalized force, 383
Generalized Fourier series, 6
Geometric boundary condition, 366
Geometrical acoustics, 405, 511
Gibb's phenomenon, 174
Gradient
for scattering, 492
in cylindrical coordinates, 672
in spherical coordinates, 665
Graphical construction
of a characteristic equation, 228
of a spatial profile, 553
of a waveform, 554
of wavefronts and rays, 408
Grazing angle, 418
Green's function, 85,481
for a cavity, 295, 296
for a rigid plane, 135
free space, 87
two-dimensional, 72, 87
Green's theorem, 118, 381
Group velocity, 202, 222, 233

## H

Half-power point, 148
Hamilton's principle, 383
Hankel function
cylindrical, 58, 268
spherical, 11, 19
Hankel transform, 146
Harmonic amplitudes, nonlinear wave, 567, 568, 603
Heavy fluid loading, 394
Helical surface wave, 61, 66, 69
Helmholtz equation, 10, 85, 217
for scattering, 480, 492
in cylindrical coordinates, 56, 267
in spherical coordinates, 2
inhomogeneous, 135
Hemisphere, in a baffle, 29, 165, 198
Hermitian, 106
Heterogeneity, 409
arbitrary dependence, 438
cylindrical region, 451
scattering from, 484, 491
vertically stratified, 418
High frequency approximation
for a spherical wave, 24
for cylindrical waves, 71
for geometrical acoustics, 407
Horizontal tangency of a ray path, 421
Horn, 187
conical, 188, 191
exponential, 193
WKB method, 207
Hyperbolic function, 226

## I

Ideal gas law, 549
Images, method of, 30, 135
Impedance boundary condition
in a cavity, 330
in a waveguide, $254,259,273$
orthogonality condition with, 254
Impedance tube, 326
Imposed boundary condition, 366
Impulsive source, 347
Incident wave, 485
Index of refraction, 418, 465
Inertia matrix, 120, 383
Infinite elements, 123
Inhomogeneous Helmholtz equation, 294
Inhomogeneous wave equation, 340, 381
Initial conditions, for a nonlinear wave, 548
Integral equation, 111
Intensity,
for a spherical harmonic series, 22
for radiation of a hemisphere, 31
from a baffled piston, 148, 166
from a spherical harmonic series, 22
from a vibrating cylinder, 70
in a ray tube, $429,437,445$
in a scattered wave, 512, 514
in a waveguide, 196, 247, 256
in the old-age stage, 595
radiated by a source, 429
using a farfield approximation, 171
Interaction of harmonics, 175, 606
Interference
in a sound beam, 154, 161, 173
in a waveguide, 260, 266
in scattering, 521
Interpolating function, 111, 116
Inverse Fourier transform, 82, 678

## J

Jacobian, 448, 646
Jump conditions, 580

## K

KHIT, see Kirchhoff-Helmholtz integral theorem
Kinematical nonlinearity, 599
Kinetic energy, 364
Kirchhoff approximation, 510
Kirchhoff-Helmholtz integral theorem
evaluation at a surface point, 90
farfield approximation, 93, 121
for a cavity, 85
for a cylinder, 94
for a half-space, 134
for a spherical cavity, 88
for radiation, 91
for scattering, 481
multipole expansion of, 99

## L

Lagrange's equations, 383
Laplace equation, 492
Laplacian
in cylindrical coordinates, 674
in spherical coordinates, 668
Legendre function, 4, 5, 8, 16, 22
Legendre polynomial, 4
Legendre series, 6, 8
Legendre's equation, 3
Leibnitz' rule, 212, 580, 583, 585
Light fluid loading, 394
Lighthill's method, see Strained coordinates
Line source, 72
finite length, 73
Linear least squares, 107, 114
Lithotripter, 657
Locally planar approximation, 408
Locally reacting
cylindrical wall, 330
effect on cavity resonance, 305
end, 323
for a waveguide, 253
for geometrical acoustics, 511

Long range communication, 656
Loudspeaker, 133, 144, 332

## M

Mass acceleration, 72, 74, 86, 105
Material derivative, 410
Material nonlinearity, 544
Matrix equation
for a cavity, 309, 350
for a waveguide, 250, 264, 266
for finite elements, 120
for radiation, 106, 113, 120
for ray tracing, 440
for Rayleigh-Ritz method, 373
Membrane shell model, 36
Modal amplitudes, see Modal series
Modal density, 346
Modal series
for a cavity, 293, 298, 300, 313, 324, 331, 340
for a circular waveguide, 243, 246, 255, 279
Modes, 336
of a cylindrical cavity, 352
of a cylindrical waveguide, 272
of a rectangular cavity, 342
of a spherical cavity, 358
of a spherical shell, 41
of a spherical waveguide, 272
two-dimensional, 217, 224, 235
Modified Bessel function, 64, 273
Momentum equation
for a heterogeneous fluid, 410
for a horn, 190
for a shock, 578
in terms of a velocity potential, 597
nonlinear wave, 544, 598
Monopole, see Green's function, Point source
Monopole amplitude, 73, 78
for scattering from a rigid body, 450, 501
Monostatic scattering measurement, 505
Morse-Penrose inverse, 107
Moving boundary condition, 612, 638
importance of, 547
Mulitvalued waveform, 565, 577
Multipath propagation, 458
Multiple scales, method of, 623
Multipole expansion, 99, 130, 496
Mutual scattering, 480

## N

Natural frequency
of a cylindrical cavity, 330, 353
of a rectangular cavity, 291, 301, 342
of a spherical cavity, 311,358
of a spherical shell, 41, 523
Neumann function
cylindrical, 58, 268, 353
spherical, 12, 15, 319
Newton's method, 563
Nondispersive wave, 202
Nonlinear propagation
in two-dimensions, 637
of a plane wave, 540
of a spherical wave, 625
Nonlinear wave equation, 598, 604, 611, 623
Nonlinearity, effects of, 568, 574, 577
Nonplanar transducer in a baffle, 163
Nonuniform plane wave, 221
Nonuniform validity, 621, 641
Normal direction
to a surface, $86,91,413$
to a vibrating plate, 638
to a wavefront, 408
Normalization
of cavity modes, 339
of transverse mode functions, 245, 255, 260, 278
Normal velocity, 133, 638
Numerical method
for a horn, 213
for a nonlinear wave, 559, 563, 649, 652
for characteristic equation, 227, 231
for nearfield of a sound beam, 158
for ray tracing equations, 440,452
for ray tube area, 448
for transport equation, 452
Nyquist criterion, 174
for a nonlinear waveform, 567

## 0

Oblique plane waves, 123, 222
Ocean, propagation in, 405
Old-age stage, 594
Omnidirectional source, 430
Order
of a Bessel function, 58
of an associated Legendre function, 16
of Neumann and Hankel functions, 11, 59
Order of magnitude, definition, 612
Orthogonality
of axial mode functions, 330
of cavity modes, 338,365
of harmonic functions, 604
of Legendre functions, 31
of transverse mode functions, 245, 255, 260, 278
Output from a horn, 188

## P

Parametric arrays, 657
Parseval's theorem, 574
Particle velocity
according to geometrical acoustics, 445
for a rigid body, 101
for a set of sources, 106
in a cylindrical wave, 62
in a nonlinear wave, 544, 548, 551
in a ray tube, 429
Perfectly transmitting boundary conditon, 123
Perturbation series, 612, 626, 639
Phase delay, due to a caustic, 450
Phase speed
for a plate, 38
for a waveguide, 223, 228, 243
in a horn, 195, 202, 204, 207, 208
in ray tracing equations, 438
nonlinear, 543, 647
Phase variable, 209, 603, 644
Phase-reversal, 556
Piecewise linear waveform, 553, 574
Piston
at the end of a cylinder, 279
in the time domain, 176
nearfield, 158
on-axis pressure, 153,159
radiated power, 171
radiation impedance, 166
Plane wave
for group velocity, 202
in a horn, 193
in a waveguide, $215,221,268$
nonlinear, 541
refraction by a cylindrical heterogeneity, 451
Plane wave approximation, 62, 123, 149
Plane wave mode, 220, 225, 261
Point source, set of, 104
Poisson solution, 540
Position
in cylindrical coordinates, 671
in spherical coordinates, 665

Potential energy, 36, 120, 364
Potential function, see Velocity potential
Power flow, 256
in a ray tube, 430,446
in a waveguide, 248, 257
Power input, 171, 197, 248, 303, 430
Power, Radiated
at an interface, 256
by a hemisphere, 32
for a spherical harmonic series, 24
from a baffled piston, 166, 171
from a vibrating cylinder, 72
in a scattered wave, 506, 512
using a farfield approximation, 172
Pressure
amplitude along a ray, 446
from the ray tracing equations, 459,497
in a cylindrical wave, 62
in a spherical cavity, 15
in terms of strained coordinates, 621, 630, 643
in terms of velocity potential, 598, 599
relation to particle velocity, see Euler's equation
scattered, see Scattered pressure
Pressure loading, of a sphere, 38, 497, 524
Pressure-release
baffle, 29
end correction, 133
wall, 226, 353
Principal curvature, 511
Product, of Fourier series, 604
Pulse, 596

## R

Radial symmetry, 23, 625
Radiation
from a piston, 157
from a spherical shell, 522
from a vibrating plate, 218, 637
from a vibrating sphere, 15,19
from an infinite cylinder, 55
Radiation damping, 45, 122, 525
Radiation impedance
of a piston, 166
of an infinite cylinder, 63
specific, 170
Radiation pressure, 655
Range determination, 506
Rankine-Hugoniot conditions, 580
Ray
definition, 408
for a nonlinear wave, 644, 647
graphical construction of, 408
in vertically stratified fluid, 417
reflected, 223, 413, 421
transmitted, 414
Ray tracing equations, 466
derivation, 438
numerical solution of, 440, 452
relation to Fermat's principle, 467
Ray tube
definition, 428, 445
Jacobian for, 448, 459
Rayleigh distance, 146, 155, 162
Rayleigh integral
application for nonplanar surface, 146, 163
derivation, 134
farfield approximation of, 138, 144
for a nonplanar transducer, 163
for a surface point, 167
for an axial field point, 153
in Cartesian coordinates, 137
in cylindrical coordinates, 144
numerical approximation of, 158
time domain version, 177
Rayleigh ratio, 364, 370
Rayleigh-Ritz method, 370
Rayleigh scattering limit, 490, 508, 517, 519
Reactive boundary, 227, 247
Reciprocity, 167, 388
Rectangular cavity, 291, 301, 342
Rectangular waveguide, 257, 298, 625
Recurrence relation
for a gamma function, 169
for Bessel functions, 270
for Legendre polynomials, 6, 24
for modified Bessel functions, 65
for second order velocity potental, 629
for spherical Bessel functions, 14, 65
Reflecting telescope, 415
Reflection
effect on shadow zones, 435
from a curved surface, 413,511
from a free surface, 421,459
in a hard-walled waveguide, 223
of waveguide modes, 298
Reflection coefficient
from a curved surface, $414,450,511$
modal, 300
Refraction
according to geometrical acoustics, 413
Renormalization, see Strained coordinate
Resonance, 296, 300, 325
effect of fluid-loading, 44, 522, 526
in a cavity, 296, 300, 322, 325, 393
of a mechanical oscillator, 44, 526
of a spherical shell, 43, 46
Retarded time, 569
Riemann solution, 541, 623
compared to nonlinear spherical wave, 632
energy conservation, 572
evaluation of, 553, 559, 564
fitting shocks to, 588
for initial conditions, 548
interpretation, 543
Rigid body, scattering from, 494, 499, 507
Rigid surface, 337, 353
Rigid walled waveguide, 218, 271
Rigid-walled cavity, 336, 353
Ring, vibrating, 145
Rodrigue's formula, 6

## S

Salinity, 412
Saturation, 595
Sawtooth waveform, 588, 592, 595
Scattering
Born approximation, 485
by a disk, 497, 501, 502
by a movable body, 499
by a sphere, 497,515
due to strong heterogeneity, 497
from a rigid body, 494
governing equations, 492
Rayleigh limit, 490
Scattering cross section, 507, 512, 517
Scattering directivity function, 519
Scattering volume, 491
Second orthogonality condition, 339
Selection of trial functions, 367
Self-refraction, 645
Separation of variables, 2, 9, 16, 53, 56, 217
Separation theorem, 374
Series
for structural displacement, 383
of forced cavity modes, 297, 312, 323
of natural cavity modes, 342
of rigid cavity modes, 381
of spherical harmonics, 17, 19
of transverse mode functions, 245, 298
of trial functions, 370
Shadow zone, 433, 435, 521
Shedding, of creeping waves, 521
Shell theory, 37

SHIE, 111, see Surface Helmholtz integral equation
Shock
condition for, 575, 646
for a sinusoidal input, 577
formation, 539, 555, 562, 564, 574
in a spherical wave, 632
in a two-dimensional wave, 651
location, 550, 577, 582, 631, 646
old age, 594
Shock formation, 545
Short wavelength behavior, 22
Side lobe, 98, 142, 147, 519
Sinc function, 141
Singular perturbation, see Strained coordinate
Slowly varying function, 603
Slowness, see Wave slowness
Small signal approximations
for linearization, 409
for nonlinear propagation, 554
for shock formation, 575
Snell's law
at an interface, 414
for vertical heterogeneity, 418, 424
from Fermat's principle, 463
Solid angle, 110, 508
Sommerfeld radiation condition
for a spherical wave, 15
for radiation, 92
for scattering, 480, 517, 522
for the Kirchhoff-Helmholtz integral, 91
from a plate, 638
in a waveguide, 218, 254
Sonar, 505
Sound beam, 142
axial pressure, 156
farfield pressure, 142, 146
intensity, $148,171,175$
nearfield, 160
nonlinear, 655
qualitative description, 162
Sound speed in a heterogenous fluid, 412, 418, 439, 492
Source
in a cavity, 291, 381, see also Point source
Source distribution, 105, 341
for surface vibration, 178
Source strength, see Monopole amplitude
Source superposition method, 104
Spatial Fourier transform, 679
Speaker cone, 148

Specular reflection, 521
Sphere
radially vibrating, $19,103,625$
scattered field, 492, 496, 513, 516, 522
Spherical aberration, 417
Spherical cap, 188
Spherical cavity, 311, 358
Spherical coordinates, 137, 665
for an elastic shell, 37, 515
for scattering, 483
Spherical Hankel function, 11, 19, 319, 516
Spherical harmonics, 36, 493
Spherical spreading, 10, 62, 92
Square wave
distortion of, 592
radiation from a piston, 172, 180
Standing wave, 223, 268, 296
State vector, 386
Stationarity
of a variational path, 463
of the Rayleigh ratio, 371
Stiffness matrix, 120, 383
Strained coordinates, 621, 642
Stringed musical instrument, 79
Structural acoustics, 1, 379
Struve functions, 169
Sturm-Liouville analysis, 244
Subsonic surface wave, 63
Supercritical incidence, 414
Superposition
of plane waves, 222, 260
of rays, 459
of standing waves, 293
Supersonic surface wave, 63, 242
Surface Helmholtz integral equation, 108, 111
derivation, 108
discretization of, 113
Symmetry
in a rectangular waveguide, 236, 266, 306
of a ray path, 422
of cavity-structure equations, 387
radial, 22, 625

## T

Table
of Fourier transform pairs, 681, 682
Tangential stiffness, 543
Target identification, 505
Target strength, 510
Taylor series, 170, 410, 465, 543, 551, 611
Temporal Fourier transform, 679

Tensor, 102
Throat impedance, 197
Time delay, 177, 443
Time domain Rayleigh integral, 177
Time, along a ray, 420, 440, 442
Trace velocity, 222
Trace wavenumber, 63, 139
Transducer impedance, 248
Transfer function, 683
Transient response, 296
in a cavity, 296
of a piston, 176
Transmission coefficient, 414
Transport equation, 444
Transverse distance, 671
Transverse mode function, 219
coupling with an elastic plate, 235
relation to forced cavity modes, 254,258 , 259 , see alsoMode function
Transverse wavenumber, 82, 353
Trial function, 365,367
Truncation
for Dowell's approximation, 394
of a Fourier series, 602
of a modal series, 280,350
of a perturbation series, 612
of a spherical harmonic series, 23,517 , 525
of equation of state, 551
Turning point, 421

## U

Ultrasonic imaging, 505, 656
Underwater acoustics, 418
Unit vector
in cylindrical coordinates, 672
in spherical coordinates, 666
Upper bound, for natural frequency, 366

## V

Variational path, 463
Velocity
in cylindrical coordinates, 675
in spherical coordinates, 669
of a shock, 581
Velocity potential, 598, 605, 623, 626
Vertical stratification, 417
Vertical tangency, of a waveform, 562, 576, 578
Vibrating cylinder, 56
Vibrating pipe, 71
Vibrating plate, 637

Virtual displacement, 384, 464
Virtual mass, 103, 502
Virtual sources, 657
Virtual work, 383
Volume velocity, 340
for multiple sources, 100, 105
of a piston, 151,166
of a source, $23,72,74,105$

## W

Water, standard properties of, 412, 551
Watson transformation, 521
Wave equation
for heterogeneous fluid, 412
nonlinear, 600
Wave packet, 204
Wave slowness, 439
Waveform, nonlinear, 539, 545, 553, 559, 633, 651
Wavefront
according to geometrical acoustics, 407, 438, 448
definition, 406
for a nonlinear wave, 644
Waveguide
circular, 267
end correction, 199
three-dimensional, 253
two-dimensional, 217
with compliant walls, 224
Waveguide mode, see Transverse mode function
Wavenumber
axial, 217
complex, 195
dispersion equation for, 202
for a cylindrical wave, 62
for a spherical shell, 37
for a surface vibration, 137
for an arbitrary surface vibration, 137
vector, 61, 137, 221
Webster horn equation, 191, 209
criteria for using, 193
derivation, 189
numerical solution of, 212
Window
for Fast Fourier Transform, 683
WKB method, 207
Woofer, 332
Work-energy principle, 364
Wraparound error, 683
Wronskian, 524, 527

## Z

Zero frequency mode, 367
Zero padding, 683
Zeros of a Bessel function, 151


[^0]:    ${ }^{1}$ M.I. Abramowitz and I.A. Stegun, Handbook of Mathematical Functions, Dover, Chap. 8 (1965).

[^1]:    ${ }^{2}$ R.D. Mindlin, "Thickness-Shear and Flexural Vibrations of Crystal Plates,"J. Appl. Phys. 22, 316-323 (1951).

[^2]:    ${ }^{3}$ It is standard practice to denote a Bessel function as $J_{\nu}(x)$, but a Neumann function often is denoted as $Y_{\nu}(x)$.

[^3]:    ${ }^{4}$ M.I. Abramowitz and I.A. Stegun, Handbook of Mathematical Functions, Dover, Chaps. 9-11 (1965).
    ${ }^{5}$ I.A. Sneddon, Special Functions of Mathematical Physics and Chemistry Oliver \& Boyd (1961).

[^4]:    ${ }^{6}$ J.O. Hirschfelder, C.F. Curtiss, and R.B. Bird, Molecular Theory of Gases and Liquids, WileyInterscience (1964).

[^5]:    ${ }^{7}$ M.C. Junger and D. Feit, Sound, Structures, and Their Interaction, Acoustical Society; Second edition, pp. 229-230 (1993 reprint).

[^6]:    ${ }^{8}$ W.E. Baker, "Axisymmetric Modes of Vibration of Thin Spherical Shells," J. Acoust. Soc. Am. 33, 1749-1758 (1961).
    ${ }^{9}$ M.C. Junger and D. Feit, ibid.

[^7]:    ${ }^{10} \mathrm{An}$ infinite value of $W_{m}$ is an artifact. If any linear system is excited at a natural frequency, the response grows in time without bound. This behavior was observed in Sect. 2.5.3. The mathematical explanation may be found in Mechanical and Structural Vibration, J.H. Ginsberg, John Wiley and Sons, Inc., Sect. 3.3 (2001).

[^8]:    ${ }^{11}$ Morse and Feshbach, Methods of Theoretical Physics, vol. 1 (1953) pp. 494-523, 655-666.

[^9]:    ${ }^{12}$ M.I. Abramowitz and I.A. Stegun, Handbook of Mathematical Functions, 9th ed., Dover, Sect. 6.3. (1965).

[^10]:    ${ }^{13}$ M.I. Abramowitz and I.A. Stegun, Handbook of Mathematical Functions, Dover, (1965) Chaps. 9-11.

[^11]:    ${ }^{14}$ Equation (9.1.23) in M.I. Abramowitz and I.A. Stegun, Handbook of Mathematical Functions, Dover, p. 360 (1965).

[^12]:    ${ }^{15}$ Pierce, ibid, p. 427.

[^13]:    ${ }^{16}$ Koopman, Song, and Fahnline, "A method for computing acoustic fields based on the principle of wave superposition," J. Acoust. Soc. Am., 86 (1989), 2433-2438.

[^14]:    ${ }^{17}$ M. Ochmann, "The full-field equations for acoustic radiation and scattering," J. Acoust. Soc. Am. 105 (1999) 2574-2584.
    ${ }^{18}$ W.H. Press, S.A. Teukolsky, W.T. Vetterling, \& B.P. Flannery, Numerical Recipes, 3rd Ed., Chap. 12, Cambridge University Press (2007).
    ${ }^{19}$ M. Ochmann, "The source simulation technique for acoustic radiation problems," Acustica 81, 512-527 (1995).

[^15]:    ${ }^{20}$ H.A. Schenck, "Improved Integral Formulation for Acoustic Radiation Problems," J. Acoust. Soc. Am. 44 (1968) pp. 41-58.

[^16]:    ${ }^{21}$ A.L. van Buren, "A Test of the Capabilities of CHIEF in the Numerical Evaluation of Acoustic Radiation from Arbitrary Surfaces," NRL Report 7160 (1970).

[^17]:    ${ }^{22}$ H.A. Schenck, ibid.

[^18]:    ${ }^{23}$ 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. A 323 (1971) 201-210.

[^19]:    ${ }^{24}$ A. Bayliss and E. Turkel, "Radiation BoundaryConditions for Wave-Like Equations," Comm. Pure Appl. Math., 33 (1980) pp.707-725.
    ${ }^{25}$ J. Assaad, J.-N. Decarpigny, C. Bruneel, R. Bossut, and B. Hamonic, "Two-dimensional radiation problems," J. Acoust. Soc. Am., 94 (1993) 562-573.
    ${ }^{26}$ D.S. Burnett, "A three-dimensional acoustic infinite element based on a prolate spheroidal multipole expansion," J. Acoust. Soc. Am., 96, (1994) 2798-2816.

[^20]:    ${ }^{27}$ ibid.
    ${ }^{28}$ R.J. Astley, G.J. Macaulay, J.-P. Coyette, and L. Cremers, "Three-dimensional wave-envelope elements of variable order for acoustic radiation and scattering. Part I. Formulation in the frequency domain," J. Acoust. Soc. Am. 103 (1998), 49-63.

[^21]:    ${ }^{1}$ J.W. Strutt Lord Rayleigh, Theory of Sound, vol. 2, 2nd ed., Dover (1945 reprint) Sect. 78.

[^22]:    ${ }^{2}$ M.I. Abramowitz and I.A. Stegun, Handbook of Mathematical Functions, Dover (1965) p. 360, Eq. (9.1.21).

[^23]:    ${ }^{3}$ I.A. Sneddon, Fourier Transforms, Dover (1995 reprint) Chap. 2.
    ${ }^{4}$ Abramowitz and Stegun, ibid., \#11.3.20.

[^24]:    ${ }^{5}$ Abramowitz and Stegun \#11.4.10.

[^25]:    ${ }^{6}$ J. Zemanek, "Beam behavior within the nearfield of a vibrating piston," J. Acoust. Soc. Am. 49 (1971) pp. 181-191.

[^26]:    ${ }^{7}$ J.W. Strutt Lord Rayleigh, Theory of Sound, vol. 2, 2nd ed., Dover (1945 reprint) Sect. 78.

[^27]:    ${ }^{8}$ M.I. Abramowitz and I.A. Stegun, Handbook of Mathematical Functions, 9th ed., Dover (1965), Eq. (9.1.21).
    ${ }^{9}$ M.I. Abramowitz and I.A. Stegun, ibid, Eqs. (12.1.3) and (12.1.7).
    ${ }^{10}$ M.I. Abramowitz and I.A. Stegun, ibid, Chap. 5.

[^28]:    ${ }^{11}$ Ping Wu, Rymantas Kazys, and Tadeusz Stepinski, "Analysis of the numerically implemented angular spectrum approach based on the evaluation of two-dimensional acoustic fields. Part I. Errors due to the discrete Fourier transform and discretization," J. Acoust. Soc. Am. vol. 99, Issue 3, pp. 1339-1348 (1996); (10 p).

[^29]:    ${ }^{12}$ D.T. Blackstock, Fundamentals of Physical Acoustics, John Wiley \& Sons (2000) pp. 461-463.
    ${ }^{13}$ A.D. Pierce, Acoustics, McGraw-Hill (1981), reprinted by ASA Press (1989) pp. 227-231.

[^30]:    ${ }^{1}$ The initials recognize G. Wentzel, H. Kramers, and L.N. Brillouin, who employed the method to analyze the Shroedinger wave equation. A detailed development is provided by C.M. Bender and S.A. Orszag, Advanced Mathematical Methods for Scientists and Engineers (1978) Chap. 10.

[^31]:    ${ }^{2}$ M.I. Abramowitz and I.A. Stegun, Handbook of Mathematical Functions, Dover, (1965) Eq. 11, pp. 485.

[^32]:    ${ }^{3}$ M.I. Abramowitz and I.A. Stegun, ibid, Eq. (11) pp. 484.

[^33]:    ${ }^{1}$ It is possible to use the forced mode formulation to determine the field of a source within an enclosed regions. The method follows the development in Sect. 6.4.2, in which a solution is constructed by adding a function $F(\bar{x})$ to the free-space source solution $G(\bar{x})$. Similarly, a natural mode formulation can be used to address boundary excitation by adding to the series a term that satisfies all boundary conditions. Equations for the modal coefficients would be obtained by requiring that the sum of the added term and the natural mode series satisfy the inhomogeneous Helmholtz equation.

[^34]:    ${ }^{2}$ M.I. Abramowitz and I.A. Stegun, Handbook of Mathematical Functions, Dover, p. 440, Eq. 10.1.47 (1965).

[^35]:    ${ }^{3}$ M.I. Abramowitz and I.A. Stegun, Handbook of Mathematical Functions, Dover, (1965) Eq. 11.3.20, p. 484.

[^36]:    ${ }^{4}$ M.I. Abramowitz and I.A. Stegun, ibid.

[^37]:    ${ }^{5}$ J.H. Ginsberg, "Derivation of a Ritz series modeling technique for acoustic cavity-structural systems based on a constrained Hamilton's principle," J. Acoust. Soc. Am. 127, 2749-2758 (2010).

[^38]:    ${ }^{6}$ E. Chaladni, Entdeckungen über die Theorie des Klanges ("Discoveries in the Theory of Sound") 1787.

[^39]:    ${ }^{7}$ M.I. Abramowitz and I.A. Stegun, Handbook of Mathematical Functions, Dover, p. 485, Eq. 11.4.5 (1965).

[^40]:    ${ }^{8}$ M.I. Abramowitz and I.A. Stegun, ibid, p. 485, Eq. 11.4.5.

[^41]:    ${ }^{9}$ J.W. Strutt Lord Rayleigh, Theory of Sound, Vol. 2, Dover (1945 reprint) Sect. 88.
    ${ }^{10}$ Dowell, E.H., and Voss, H.M. (1962). "The effect of a cavity on panel vibration," AIAA J., 1, 476-477.

    Dowell, E.H., Gorman, G.F., and Smith, D.A. (1977). "Acoustoelasticity: General theory, acoustic natural modes and forced response to sinusoidal excitation, including comparisons with experiment" J. Sound Vib., 52, 519-542.

[^42]:    ${ }^{11}$ Weinstock, Calculus of Variations, Dover (1974).

[^43]:    ${ }^{12}$ Rayleigh is widely recognized for this formulation because it appears in his monumental texts on The Theory of Sound. Walther Ritz was a pioneering nuclear physicist who independently developed

[^44]:    (Footnote 12 continued)
    the formulation. He came to it by applying variational calculus concepts he first developed for quantum mechanics. The Rayleigh-Ritz formulation actually is a special application of Ritz' general approach, which is known as the Ritz series method, see J.H. Ginsberg, Mechanical and Structural Vibrations, John Wiley and Sons (2000), Chap.6. This approach has been used for many analyses in physics and engineering. It also is the foundation for the finite element method. Thus, it is remarkable that Ritz wrote only two papers that discussed acoustic-type systems: "On the new method for solving some variational problems of mathematical physics" (1908) and "Theory of the transverse oscillations of a square plate with free boundaries" (1909). These were the first to explain mathematically the occurrence of Chaladni lines in plate vibrations.
    ${ }^{13}$ L. Meirovitch, Principles and Techniques of Vibrations, Prentice-Hall (1997) Chap. 8.

[^45]:    ${ }^{14}$ J.H. Ginsberg, Engineering Dynamics, Cambridge University Press (2008) Chap. 7

[^46]:    ${ }^{15}$ Ginsberg, Mechanical and Structural Vibrations (2001), Ch. 11.

[^47]:    ${ }^{1}$ L.M. Brekohovskikh and Yu. P. Lysanov, Fundamentals of Ocean Acoustics, Springer-Verlag, 1982.
    ${ }^{2}$ A.D. Pierce, Acoustics: An Introduction to Its Physical Principles and Applications, Acoustical Society of Amer (1989).
    ${ }^{3}$ F.B. Jensen, W.A. Kuperman, M.B. Porter, and H. Schmidt, Computational Ocean Acoustics, 2nd ed (2011).

[^48]:    ${ }^{4}$ This assertion that the set of wavefronts at a specific instant may be identified by following the evolution of a specific wavefront is only valid in the context of linear acoustics. For example, it assumes that the signal does not alter the properties of the fluid, even though the speed of sound depends on the pressure. If nonlinear effects are included, wavefronts and rays might also depend on the pressure and particle velocity associated with a specific phase. This condition will be encountered in Sect. 13.4.2.

[^49]:    ${ }^{5}$ A.D. Pierce, Acoustics, McGraw-Hill (1981), reprinted by ASA Press (1989) pp. 33-34.

[^50]:    ${ }^{6}$ L. Brekohovskikh and Yu. Lysanov, Fundamentals of Ocean Acoustics, Springer-Verlag (1982).

[^51]:    ${ }^{7}$ Pierce, ibid.

[^52]:    ${ }^{8}$ W. Munk, "Sound channel in an exponentially stratified ocean, with application to SOFAR," Acoustical Society of America, Vol. 55 (1974) pp. 220-226.

[^53]:    ${ }^{9}$ F.B. Jensen et al., W.A. Kuperman, M B. Porter, and H. Schmidt, Computational Ocean Acoustics, 2nd ed., Chap. 3, Springer (2011).

[^54]:    ${ }^{10}$ A.D. Pierce, Acoustics, Sect. 9-4, ASA Books reprint of 1981 edition.

[^55]:    ${ }^{11}$ O. Bühler, Waves and Mean Flows, Cambridge University Press (2009) pp. 81-83.

[^56]:    ${ }^{1}$ M.I. Abramowitz and I.A. Stegun, Handbook of Mathematical Functions, Dover, (1965), p. 360 for Eq. (9.1.210), and p. 484 for Eq. (11.3.20).

[^57]:    ${ }^{2}$ J.W. Strutt, Lord Rayleigh, "On the light from the sky, its polarization and colour," Philosophical Magazine, series 4, vol. 41 (1871) pp. 107-120.

[^58]:    ${ }^{3}$ A.D. Pierce, Acoustics, Mc-Graw-Hill (1981), ASA reprint, (1989), Sect. 8-8.

[^59]:    ${ }^{4}$ C.M. Bender \& S.A. Orszag, Advanced Mathematical Methods for Scientists and Engineers (1978).

[^60]:    ${ }^{5}$ M.I. Abramowitz and I.A. Stegun, Handbook of Mathematical Functions, Dover, (1965), p. 440, Eq. 10.1.47.

[^61]:    ${ }^{6} \mathrm{H}$. Stenzel, "On the perturbation of a sound field caused by a rigid sphere," Electr. Nachrichtentech., vol. 15, (1938) pp. 71-28.

[^62]:    ${ }^{7}$ G.N. Watson, "The Diffraction of Electric Waves by the Earth; the Transmission of Electric Waves Around the Earth," Proc. Royal Soc. London A95 (1918), pp. 83-88, 546-563.

[^63]:    ${ }^{8}$ M.C. Junger and D. Feit, "Sound, Structures, and Their Interaction," 2nd ed., MIT Press, ASA reprint, Chap. 12 (1986).
    ${ }^{9}$ A.D. Pierce, Acoustics, Mc-Graw-Hill (1981), ASA reprint, (1989), Chap. 9.

[^64]:    ${ }^{10}$ J.B. Keller, "Geometric Theory of Diffraction," J. Opt. Soc. Am.52, (1962) pp. 116-130.

[^65]:    ${ }^{11}$ M.I. Abramowitz and I.A. Stegun, ibid., p. 439. Use Eq. (10) to eliminate $n_{n-1}$ (z) and $j_{n-1}(z)$ in Eq. 10.1.31.

[^66]:    ${ }^{1}$ B. Riemann, "Über die Fortpflanzung ebener Luftwellen von endlicher Schwingungsweite," Abhandlungen der Königlichen Gesellschaft der Wissenschaften zu Göttingen, Vol. 8 (1859) pp. 245-264.
    ${ }^{2}$ S. Earnshaw. "On the mathematical theory of sound," Trans. Royal Soc. London (1860) pp. 133148.
    ${ }^{3}$ J.W. Strutt, Lord Rayleigh, "Theory of Sound," Vol 2, (1877), Dover reprint (1945) Articles 252 and 253.
    ${ }^{4}$ Lamb, Hydrodynamics, (1916), Dover reprint, (1945). Articles 260 and 261.
    ${ }^{5}$ S.D. Poisson, "Memoire sr lat théorie de son," J. l'École Polytechnique Paris, Vol 7 (1808) pp. 319-392.

[^67]:    ${ }^{6}$ S.D. Poisson, " Memoire sr la théorie de son," J. l'École Polytechnique Paris, Vol. 7 (1808) pp. 319-392.

[^68]:    ${ }^{7}$ R.T. Beyer, "The Parameter B/A," M.F. Hamilton \& D.T. Blackstock, Nonlinear Acoustics, eds., Acoustical Society of America (2008) p. 34.

[^69]:    ${ }^{8}$ M.I. Abramowitz and I.A. Stegun, Handbook of Mathematical Functions, Dover (1965) Eq. (9.1.21).

[^70]:    ${ }^{9}$ I.E. Fubini-Ghiron. "Anomalies in acoustic wave propagation of large amplitudes," Alta Freq. Vol. 4 (1935) pp. 530-581.

[^71]:    ${ }^{10} \mathrm{~A} 200 \mathrm{~dB}$ signal in water has an amplitude that is 1.4 atm . Hence, generating such a signal without causing cavitation requires that the depth be sufficiently large that the hydrostatic pressure exceeds the amplitude.

[^72]:    ${ }^{11}$ Pierce, ibid., pp. 576-577.

[^73]:    ${ }^{12}$ R.L. Bisplinghoff, H. Ashley, and R.L. Halfman,Aeroelasticity, Addison Wesley (1955) Chap. 3
    ${ }^{13}$ S. Goldstein, Lectures in Fluid Mechanics, Wiley-Interscience (1960) Chap. 4
    ${ }^{14}$ D.T. Blackstock, Physical Acoustics, Wiley (2000) pp. 84-86.

[^74]:    ${ }^{15}$ M.F. Hamilton and C.L. Morfey, "Model Equations," M.F. Hamilton and D.T. Blackstock, Nonlinear Acoustics, eds., Acoustical Society of America (2008) Chap. 3.

[^75]:    ${ }^{16}$ M.J. Lighthill, "A technique for rendering approximate solutions to physical problems uniformly valid," Philosophical Mag., Vol. 40 (1949) 1179-1201.

[^76]:    ${ }^{17}$ A.H. Nayfeh and A. Kluwick, "A comparison of three perturbation methods for nonlinear waves," J. Sound and Vibration (1976) pp. 293-299.

[^77]:    ${ }^{18}$ M.F. Hamilton \& D.T. Blackstock, Nonlinear Acoustics, eds., Acoustical Society of America (2008).

[^78]:    ${ }^{19}$ D.T. Blackstock, "On plane, spherical, and cylindrical waves of finite amplitude in lossless fluids," J. Acoust. Soc. Am., Vol. 36 (1966) pp. 217-219.

[^79]:    ${ }^{20}$ H.C. Miao and J.H. Ginsberg, "Finite amplitude distortion and dispersion of a nonplanar mode in a waveguide," J. Acoust. Soc. Amer., Vol. 80 (1986), pp. (911-920).

[^80]:    ${ }^{21}$ J.H. Ginsberg, "A new viewpoint for the two-dimensional nonlinear acoustic wave radiating from a harmonically vibrating flat plate," J. Sound and Vib., Vol. 63 (1978) pp. 151-154. The analysis of the plate problem by J.H. Ginsberg, "Perturbation Methods," M.F. Hamilton and D.T. Blackstock, eds., Nonlinear Acoustics, Acoustical Society of America (2008) Chap. 10 repeated the misinterpretation.

[^81]:    ${ }^{22}$ M.F. Hamilton and D.T. Blackstock, ibid.

[^82]:    ${ }^{1}$ I.N. Sneddon, Fourier Transforms, Dover reprint (1951).
    H.F. Davis, Fourier Series and Orthogonal Functions, Dover reprint (1963).

[^83]:    ${ }^{2}$ A. Erdelyi Bateman Project Manuscript, Vol. I - Table of Integral Transforms (1954) Ch. III.

