

Stephen Nettel

# Wave Physics

Oscillations – Solitons – Chaos

Fourth Edition

# Wave Physics

Stephen Nettel

# Wave Physics

Oscillations – Solitons – Chaos

Fourth Edition

With 67 Figures and more than 100 Problems  
with Hints for Solution and Numerous Examples



Springer

Professor Stephen Nettel  
Rensselaer Polytechnic Institute  
Department of Physics  
Troy, NY 12180-3590, USA

*Contributors to Chapter 8:*

Professor Dr. Andrei V. Gaponov-Grekov  
Professor Dr. Mikhail I. Rabinovich  
Institute of Applied Physics  
Russian Academy of Sciences  
ul. Ulyanova 46  
603 600 Nizhnii Novgorod, Russia

Professor Dr. Martin Gutzwiller  
IBM T.J. Watson Research Center  
P.O. Box 218  
Yorktown Heights, NY 10598, USA

ISBN 978-3-540-87907-7

e-ISBN 978-3-540-87908-4

DOI 10.1007/978-3-540-87908-4

Library of Congress Control Number: 2008940575

© 2009, 2003, 1995, 1992 Springer-Verlag Berlin Heidelberg

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

The use of general descriptive names, registered names, trademarks, 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.

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

[springer.com](http://springer.com)

## Preface to the Third Edition

This is a text for the third semester of undergraduate physics for students in accelerated programs, who typically are preparing for advanced degrees in science or engineering. The third semester is often the only opportunity for physics departments to present to students who are not physics majors a coherent background in the physics of waves, required later for confident handling of applied problems, especially applications based on quantum mechanics.

Physics is a coherent subject. It is often found that the going gets easier as one goes deeper, learning the mathematical connections tying together the various phenomena. Even so, the steps that took us from classical wave physics to Heisenberg's "Physical Principles of Quantum Theory" were, as a matter of history, harder to take than later steps dealing with detailed applications. With these considerations in mind, the classical physics of oscillations and waves is developed here at a more advanced mathematical level than is customary in second year courses. This is done to explain the classical phenomena, but also to provide background for the introductory wave mechanics, leading to a logical integration of the latter subject into the presentation. Concurrently, detailed applications of quantum mechanics are beyond the mission of the text. The concluding chapters on nonlinear waves, solitons, and chaos broaden the previously established concepts of wave behavior, while introducing the reader to important topics in current wave physics.

Experience teaching the course has led in this third edition to the author's fully presenting material dealing with Green's functions, rather than trying to save course time by burying the material in problems and an appendix, as heretofore. This extension has led, in turn, to a broadening of Chap. 6 on wave mechanics, with a treatment of the Green's function for harmonic oscillators. In this way, along with additions to examples and problems, the integration of wave mechanics with classical wave physics has been reinforced.

The text begins with a full chapter of mathematics that develops Fourier analysis in the context of generalized functions. The idea is to expose students to the new concepts while there is yet a minimum of "pressure", and then allow them to assimilate the concepts through the chapters on classical physics. In this way, when it comes to the wave mechanics in Chap. 6, students should

be able to distinguish the new physics from mathematics. M. J. Lighthill, in his monograph “Fourier Analysis and Generalized Functions”, says that “The theory of generalized functions ... greatly curtails the labor of understanding Fourier transforms”. Application of generalized functions (Schwartz’ “théorie des distributions”) avoids undefined integrals like  $\int_{-\infty}^{\infty} e^{ikx} dk$  still prevalent in texts on quantum mechanics. Chapter 2, on oscillations, may seem somewhat in the nature of a review. The physics material thus starts at the point where the first exposure in the freshman year has been made, but may have led to only a partial grasp.

It is hoped that this book will be useful to various categories of readers, honors math and physics students, graduate students studying for Ph.D. qualifiers, as well as professors looking for examination problems. Problems displace tedium by presenting a challenge. Here many are used to establish significant results. This makes for meaningful activity, as well as for efficiency in the learning process.

The prerequisite for the text is limited to a standard calculus based freshman course in mechanics and electricity and magnetism. The text is an outgrowth of the Honors Physics III course which the writer had been frequently teaching at Rensselaer since 1973. He believes that some selection according to the purpose of the course will need to be excised. Only in this way will the first six chapters of this extended edition fit comfortably within a standard fourteen week semester. Selection is facilitated by the fact that the more advanced topics occur towards the end of each chapter.

Alternatively, following a suggestion of Professor Lyle Roelofs<sup>1</sup>, physics departments might consider offering Chap. 1, Mathematical Foundations, in a mini semester, if there is one available.

Saarbrücken,  
June 2002

*Stephen Nettel*

---

<sup>1</sup> L. Roelofs, Book Reviews, D.J. Griffiths Ed., Am. J. Phys. **69**, 922 (2001).

## Preface to the Second Edition

A number of examples and problems to elucidate basic concepts have been added, and typographic errors corrected.

The first edition has now been used a number of times at Rensselaer in second year courses using the interactive method of teaching. This method includes regular problem-solving sessions where students work together in groups with aid from special work sheets. There is input from more senior students, graduate and undergraduate, acting as tutors. It was discovered that with this method *Wave Physics* can be used by a wider selection of individuals to advantage than the honors students for whom the text was originally intended. The main factor in a student's success appeared to be the quality of his or her mathematical preparation. It is a pleasure to thank the many students who participated as tutors. Special thanks go to Howard Goldowsky, Byong Kim, and Richelle Thompson who carried much of the responsibility over the various classes. Our teaching experience has influenced the present revision.

Troy,  
August 1994

*Stephen Nettel*

# Preface to the First Edition

This is a text for the third semester of undergraduate physics for students in accelerated programs who typically are preparing for advanced degrees in science or engineering. The third semester is often the only opportunity for physics departments to present to students who are not physics majors a coherent background in the physics of waves required later for confident handling of applied problems, especially applications based on quantum mechanics.

Physics is an integrated subject. It is often found that the going gets easier as one goes deeper, learning the mathematical connections tying together the various phenomena. Even so, the steps that took us from classical wave physics to Heisenberg's "Physical Principles of Quantum Theory" were, as a matter of history, harder to take than later steps dealing with detailed applications. With these considerations in mind, the classical physics of oscillations and waves is developed here at a more advanced mathematical level than is customary in second year courses. This is done to explain the classical phenomena, but also to provide background for the introductory wave mechanics, leading to a logical integration of the latter subject into the presentation. The concluding chapters on nonlinear waves, solitons, and chaos broaden the previously established concepts of wave behavior, while introducing the reader to important topics in current wave physics.

The text begins with a full chapter of mathematics that develops Fourier analysis in the context of generalized functions. The idea is to expose students to the new concepts while there is yet a minimum of "pressure", and then allow them to assimilate the concepts through the chapters on classical physics. In this way, when it comes to the wave mechanics in Chap. 6, students should be able to distinguish the new physics from mathematics. M. J. Lighthill, in his monograph "Fourier Analysis and Generalized Functions", says that "The theory of generalized functions ... greatly curtails the labor of understanding Fourier transforms". Application of generalized functions (Schwartz' "théorie des distributions") avoids undefined integrals like  $\int_{-\infty}^{\infty} e^{ikx} dk$  still prevalent in texts on quantum mechanics. Chapter 2, on oscillations, may seem somewhat in the nature of a review. The physics material thus starts at the point where the first exposure in the freshman year has been made, but may have led to only a partial grasp.

It is hoped that this book will be useful to various categories of readers, honors math and physics students, graduate students studying for Ph.D. qualifiers, as well as professors looking for examination problems. Problems displace tedium by presenting a challenge. Here many are used to establish significant results. This makes for meaningful activity, as well as for efficiency in the learning process.

The prerequisite for the text is limited to a standard calculus based freshman course in mechanics and electricity and magnetism. The text is an outgrowth of the Honors Physics III course which the writer has been frequently teaching at Rensselaer since 1973. He has found that the first six chapters can be taught comfortably within a standard fourteen week semester.

Troy,  
February 1992

*Stephen Nettel*

## Acknowledgements

It is a pleasure to express my gratitude to Dr. Ernst F. Hefter, the original editor at Springer-Verlag, whose interest has made this text a reality. I have had the benefit of some hitherto unpublished work of Dr. Hefter's on solitons, contributing to the presentation of nonlinear waves in Chap. 7. Thanks are likewise due to Dr. Natasha Aristov of Springer-Verlag for her critical reading. The writer is much indebted to Dr. Hans J. Kölsch, the present editor, for his advice and encouragement during the preparation of this, the third, edition.

The contributions of Professor Martin C. Gutzwiller and of Professors A. V. Gaponov-Grekhov and M. I. Rabinovich are deeply appreciated.

Professor G.B. Whitham of Caltech was kind enough to send me some unpublished notes on inverse scattering in soliton theory. Professor P. Banderet and H. Beck at the University of Neuchâtel, in Switzerland, were generous with various help. The "Further Reading" at the end of each chapter is an indication of sources.

The honors students at Rensselaer Polytechnic Institute who have taken the course which led to this text have contributed by their scientific curiosity as by their tolerance. Later, it was discovered that *Wave Physics* can be used to advantage by a wider selection of individuals than honors students when using an interactive method of teaching. Special thanks go to Howard Goldowsky, Byong Kim, and Richelle Thompson who carried much of the responsibility for the various classes.

The idea of writing a text came at the time when I had been initiated into Transcendental Meditation, which imparted the desire to add to my professional activity in some creative way. Lastly, I express gratitude to Irmgard Blandfort for her moral support, and for taming our computer from time to time.

# Contents

<b>1. Mathematical Foundation for Wave Physics . . . . .</b>	<b>1</b>
1.1 Introduction to the Concepts . . . . .	1
1.2 Important Integrals . . . . .	5
1.3 Complex Numbers . . . . .	7
1.4 The Dirac Delta Function . . . . .	9
1.5 Fourier Analysis . . . . .	11
1.6 Hilbert Space . . . . .	17
Problems . . . . .	24
Further Reading . . . . .	31
<b>2. Oscillations of Mechanical and Electrical Systems . . . . .</b>	<b>33</b>
2.1 The Systems and Their Equations . . . . .	33
2.2 Natural Motion of Systems . . . . .	36
2.3 Forced Motion . . . . .	41
2.4 Forced Motion and the Green's Function . . . . .	44
2.5 Oscillators . . . . .	46
2.6 Summary . . . . .	48
Problems . . . . .	49
Further Reading . . . . .	53
<b>3. Waves on Stretched Strings . . . . .</b>	<b>55</b>
3.1 Equation of Motion of a String . . . . .	55
3.2 Natural Motion of a String . . . . .	58
3.3 The Normal Modes . . . . .	60
3.4 Forced Motion of a Stretched String . . . . .	67
3.5 Green's Functions for a Stretched String . . . . .	71
Problems . . . . .	78
Further Reading . . . . .	85
<b>4. Electromagnetic Waves . . . . .</b>	<b>87</b>
4.1 Maxwell's Equations in Integral Form . . . . .	87
4.2 Maxwell's Equations in Differential Form . . . . .	88
4.3 Plane Electromagnetic Waves in Free Space . . . . .	92
4.4 Distributed Electromagnetic Systems – Cavities . . . . .	98

4.5	The Vector Potential and Related Solutions of Maxwell's Equations . . . . .	103
4.6	Dipole Radiation . . . . .	107
4.7	Electromagnetism and the Green's Function . . . . .	110
	Problems . . . . .	117
4.A	Appendix . . . . .	124
	Further Reading . . . . .	127
5.	<b>Light – Physical Optics, Refraction</b> . . . . .	129
5.1	The Nature and Generation of Light . . . . .	129
5.2	Diffraction . . . . .	133
5.3	X-ray Diffraction . . . . .	139
5.4	EM Waves in Dielectrics. Refraction . . . . .	141
5.5	The Magic Rule in Three Dimensions . . . . .	147
	Problems . . . . .	151
	Further Reading . . . . .	156
6.	<b>Wave Mechanics</b> . . . . .	157
6.1	Origin of Schroedinger's Wave Equation . . . . .	157
6.2	Postulates of Wave Mechanics . . . . .	161
6.3	Motion of a Free Particle. The Heisenberg Uncertainty Principle . . . . .	171
6.4	Wave–Particle Duality and Loss of Determinism . . . . .	179
6.5	Driven Motion and the Green's Function for the Harmonic Oscillator . . . . .	181
6.6	Scope of Quantum Mechanics . . . . .	191
	Problems . . . . .	206
	Further Reading . . . . .	215
7.	<b>Nonlinear Waves on Water – Solitons</b> . . . . .	217
7.1	Linear Surface Waves on Water . . . . .	218
7.2	Dispersion. Group Velocity . . . . .	221
7.3	Nonlinear Waves . . . . .	223
7.4	Solitons . . . . .	229
7.5	Inverse Scattering . . . . .	233
	Problems . . . . .	241
	Further Reading . . . . .	246
8.	<b>Nonlinear Phenomena – Chaos</b> . . . . .	249
8.1	Nonlinear Physics – Chaos and Order . . . . .	250
	Further Reading . . . . .	264
8.2	Quantum Chaos . . . . .	265
	<b>Hints for Solution</b> . . . . .	277
	<b>Index</b> . . . . .	285

# 1. Mathematical Foundation for Wave Physics

**Abstract.** The chapter begins with an introduction to the main concepts in generalized functions theory. The development of the mathematics of wave physics itself begins by evaluating certain essential integrals involving gaussians. A brief section on complex variables yields additional integrals. After its intuitive introduction, the Dirac delta function is shown to be an example of a generalized function, that is, it is given by a certain family of equivalent sequences. The members of such sequences must be good functions, for example, gaussians. The designations “equivalent”, “good” are carefully defined. The Fourier representation as a decomposition of functions of spatial variables into harmonic functions of all possible wavelengths is introduced. Using the analogy of basis vectors in three dimensions, the idea is developed that harmonic waves form a complete basis in an infinite dimensional function space (Hilbert space). The formula for the Fourier transform is obtained in analogy to projecting out vector components, the Dirac delta function serving as the analogue of the Kronecker delta. As summary a comparison is given between formulae governing representation in vector and Hilbert space of arbitrary vectors and functions, respectively. Worked examples include the Fourier transform of a square wave, the computer synthesis of the square wave from its transform, and orthonormality of the Dirac delta functions. Problem 1.17, which develops the Green’s function for diffusion, is particularly recommended as background for Chap. 6, “Wave Mechanics”.

## 1.1 Introduction to the Concepts

Waves play a pervasive role in nature. There are the mechanical waves, including seismic waves, sound waves in air, water waves. There are electromagnetic waves, and underlying all matter, quantum mechanical waves. These diverse wave phenomena are understood on the basis of a few unifying mathematical concepts.

In the study of particle mechanics we use vectors to describe displacements, velocities, and other quantities of motion. In three dimensions any vector  $\mathbf{F}$  can be constructed from three building blocks, the standard unit vectors  $\mathbf{i}, \mathbf{j}, \mathbf{k}$ :

$$\mathbf{F} = A_1 \mathbf{i} + A_2 \mathbf{j} + A_3 \mathbf{k} ,$$

$A_1, A_2, A_3$  being the component amplitudes which determine  $\mathbf{F}$ . The corresponding building blocks for waves are, in one dimension,  $\sin kx, \cos kx$ , if we are working in space and  $\sin \omega t, \cos \omega t$  in the time domain. Here  $k$  is called the wave number. It is related to the wavelength  $\lambda$  by

$$k = \frac{2\pi}{\lambda}.$$

Similarly,  $\omega$ , the angular frequency in radians/seconds, is related to the period  $T$  in seconds, or the frequency  $\nu$  in Hz (cycles/s) by

$$\omega = 2\pi\nu = \frac{2\pi}{T}.$$

For mathematical convenience we shall often combine the harmonic functions into the complex exponential:

$$e^{ikx} = \cos kx + i \sin kx \quad i = \sqrt{-1}.$$

We say that a particle in three dimensions has three degrees of freedom. The quantities of motion for continuous systems such as a long string under tension have infinitely many degrees of freedom. There are an infinite number of points along the line for each of which the motion must be specified. From a different perspective, the displacement of the string is often analyzed as a continuous function of position at any given time. We should not be surprised to learn that such continuous functions are conveniently represented by vectors in a “function space” which has an infinite number of dimensions, thereby allowing for an infinite number of degrees of freedom. This function space with an infinite number of building blocks or *basis functions* is a key concept in the physics of waves. What one wants to know is whether the basis functions form a *complete set*. By complete set we shall mean that the physical functions of interest can indeed, be adequately described by appropriate sums or integrals over the available set of basis functions. The question of completeness as required in theoretical physics was studied by the German mathematician David Hilbert (1862–1943). In Sect. 1.6 we shall review these ideas. Here the necessary infinite number of building blocks or basis functions is obtained by noting that  $k$  (or  $\omega$ ) is a continuous variable taking on any one of the infinite possible values that lie in the range  $\infty > k > -\infty$ . Each new positive value of  $k$  yields two new basis functions  $\sin kx, \cos kx$ .

Going to the time domain we find that the idea of  $\sin \omega t, \cos \omega t$  as basic components is not new to us. Anyone interested in high-fidelity reproduction of music is familiar with the technique of decomposing the signal  $F(t)$  into separate signals, each with a single frequency  $\omega$ , a definite amplitude  $B(\omega)$ , and a phase  $\delta\omega$ .  $F(t)$  is then given as the sum of all these separate signals, that is, as an integral over all the possible frequencies:

$$F(t) \sim \int_0^{\infty} B(\omega) \sin(\omega t + \delta\omega) d\omega = \int_{-\infty}^{\infty} A(\omega) e^{i\omega t} d\omega$$

where in the last integral  $A(\omega)$  is a complex number chosen so as to take into account the phase shift  $\delta\omega$  and to make  $F(t)$  real valued. The ideal high-fidelity equipment is that which processes (converts and amplifies) all the different frequency components in the same way.

The possibility of resolving functions of position into harmonic components of all possible wavelengths was introduced by the French mathematician Joseph Fourier in his “Theorie analytique de la chaleur” in 1822. In one dimensional space, we write

$$F(x) \sim \int_{-\infty}^{\infty} A(k)e^{ikx} dk.$$

In nature we find pairs of “opposites” or antipodes, point and infinity, etc. We shall learn that the exponential wave  $\exp(ikx)$  also has its opposite, the so-called *Dirac delta function*  $\delta(x)$  [or  $\delta(k)$ ] named after the British Nobel prize winning physicist P.A.M. Dirac. The function  $\delta(x)$  has an infinitely high and narrow peak at  $x = 0$ . We have for  $\delta(x)$  the properties:

$$\begin{aligned}\delta(x) &= 0 & x \neq 0 \\ \int_{-\infty}^{\infty} \delta(x) dx &= 1 \\ \int_{-\infty}^{\infty} F(x)\delta(x) dx &= F(0).\end{aligned}$$

The last property, true for any “physical” function  $F(x)$ , follows intuitively from the localization of  $\delta(x)$  at  $x = 0$ . If  $\exp(ikx)$  represents some infinitely extended phenomenon to be found anywhere between  $\infty > x > -\infty$ ,  $\delta(x)$  is its antipode representing total localization. We shall see the need for a delta function as we develop the mathematics of wave physics, particularly in Chap. 6, which deals with matter waves (quantum mechanics).

We note that, taken together, all the amplitudes of the wave components  $A(k)$  determine any particular function  $F(x)$  which is being Fourier analyzed. The central problem of Fourier analysis is, given a function  $F(x)$ , how can one find these (complex) expansion coefficients (amplitudes and phases)? This is called finding the *Fourier transform* of the given function. The procedure for obtaining the transform is analogous to projecting out of any given vector one of its components. For vectors this depends on the *orthonormality* of the unit vectors  $i, j, k$ . For functions, we shall see that  $u_k \equiv 1/\sqrt{2\pi} \exp(ikx)$  may be considered normalized, and two basis functions  $u_k, u_{k'}$  with  $k \neq k'$  orthogonal.

Orthonormal for  $i, j, k$  means that these vectors are perpendicular to each other (orthogonal) and each vector has length unity (normal). The procedure to establish the orthonormality of functions analogous to taking the vector

dot product is to take appropriate “overlap integrals”, see Problem 1.20. The overlap integrals are obtained by integrating the product of the complex conjugate of the first function  $u_k^*$  and of the second function  $u_{k'}$  over the range  $\infty > x > -\infty$ :

$$\int_{-\infty}^{\infty} u_k^*(x) u_{k'}(x) dx = 1/2\pi \int_{-\infty}^{\infty} e^{-i(k-k')x} dx .$$

It was suggested by Dirac that in physical calculations the overlap integrals should yield the delta function  $\delta(k - k')$ . Since if  $k \neq k'$ ,  $\delta(k - k') = 0$ , functions with different  $k$  will have zero overlap, and if  $k = k'$ ,  $u_k(x)$  would in some sense be normalized. Thus, we see that when seeking Fourier transforms, that is, when projecting the individual wave components out from a function we shall need to understand the Dirac delta function.

This discussion needs some mathematical justification. As it stands, the overlap integral of the two functions  $u_k, u_{k'}$  is not yet defined, that is, no limit is approached as the range of integration is extended to infinity. In addition, the Dirac delta function is not a function in the sense that we could ascribe a value to it at every value of the independent variable  $x$ . These difficulties have a physical origin. If  $\exp(ikx)/\sqrt{2\pi}$  represents a wave corresponding to some phenomenon, then usually in physical problems that wave must be presumed to have a beginning (in space, or time if we replace  $x$  by  $t$ ), and an end. It turns out that if we use the more rigorous *theory of generalized functions* to handle the delta function, and proceed in a consistent fashion from there, then the effect of distant yet physical boundaries is automatically supplied and the lack of mathematical definition of the orthonormalizing integrals disappears.

The theory of generalized functions begins with the identification of certain functions as *good functions*. These, as we shall see in the next section, are functions that go to zero in a particularly regular way as  $x$  goes to  $\pm$  infinity. The definition of generalized functions has its origin in the observation that whereas  $\delta(x)$  is not a function in the ordinary sense, it can be simulated by certain infinite sequences of good functions. A generalized function  $f(x)$  is defined through an infinite sequence of integrals by the equation

$$\int_{-\infty}^{\infty} f(x) F(x) dx \equiv \lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} f_n(x) F(x) dx .$$

Here  $f_n(x)$  is any sequence of good functions  $f_n(x)$ ,  $n = 1, 2, 3, \dots$  which give a desired limiting value to the integral, for example,  $F(0)$  for  $f(x) = \delta(x)$ .  $F(x)$  is, again, any good function, and can be thought of as a “test function” for various sequences. In the case of the delta function the members of each sequence will become progressively sharper and higher as  $n$  increases. It is also possible to define Fourier transforms (FT) of generalized functions by using sequences whose members are FTs of the original sequences’ members.

Once the generalized function has become familiar through repeated application, its abstract nature will, just like that of the ordinary derivative, not present any hindrance. The theory of generalized functions has its origin in Laurent Schwartz's "Theorie des distributions" which appeared in 1951, and was developed in the form outlined here by George Temple and M.J. Lighthill. We list under Further Reading at the end of the chapter the monograph by M.J. Lighthill. The theory allows one to include in Fourier analysis not only the delta function and its derivatives, but quite simple functions such as constants and powers of  $x$ , which were formerly excluded. Our approach to Fourier analysis in terms of an orthonormal basis and generalized functions is expedient both for learning the analysis, as well as for developing the physics of classical and matter waves.

## 1.2 Important Integrals

The function  $g(x)$  is defined as

$$g_a(x) \equiv \sqrt{a/\pi} e^{-ax^2}, \quad (1.1)$$

$a$  being a positive real number. This function, named after the German mathematician Friederich Gauss (1777–1855), is used extensively not only in physics but also in various statistical analyses. We calculate here a number of integrals with integrands containing gaussians, which occur again and again in all branches of physics.

The limits of these integrals are at  $x = \pm\infty$ . Because of the rapid approach of  $g_a(x)$  to 0 as  $|x| \rightarrow \infty$  the integrals will be well defined. Actually,  $g_a(x)$  will be an important example of what we have already called a *good function*. The formal definition of a good function is one which is everywhere differentiable any number of times, and it and all its derivatives are on the order of at most  $|x|^{-N}$  as  $|x| \rightarrow \infty$  for all  $N$ .

The first integral of importance to us is

$$I_1 = \int_{-\infty}^{\infty} e^{-ax^2} dx, \quad (1.2)$$

which we can find by using a trick. We write

$$\begin{aligned} I_1^2 &= \int_{-\infty}^{\infty} e^{-ax^2} dx \int_{-\infty}^{\infty} e^{-ay^2} dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-a(x^2+y^2)} dx dy. \end{aligned} \quad (1.3)$$

We convert to polar coordinates  $r, \theta$ , remembering that  $dxdy$  becomes  $rdrd\theta$ .  $I_1^2$  becomes

$$I_1^2 = \int_0^\infty r dr \int_0^{2\pi} d\theta e^{-ar^2} = 2\pi \int_0^\infty e^{-ar^2} r dr = \frac{\pi}{a}. \quad (1.4)$$

The desired result is

$$\int_{-\infty}^\infty e^{-ax^2} dx = \sqrt{\frac{\pi}{a}}. \quad (1.5)$$

In the introduction we have suggested the convenience of “normalizing” functions, that is, of adjusting their relative magnitudes in some definite way. From (1.3) we see that  $g_a(n)$  has been defined so that

$$\int_{-\infty}^\infty g_a(x) dx = 1. \quad (1.6)$$

The next integral of interest is

$$I_2 = \int_{-\infty}^\infty \cos(kx) e^{-ax^2} dx. \quad (1.7)$$

What trick can we use? Let us expand  $\cos(kx)$  in a Taylor series:

$$\cos kx = 1 - (kx)^2/2! + (kx)^4/4! - \dots. \quad (1.8)$$

Next, we differentiate  $I_1$  under the integral sign with respect to  $a$  and find

$$\int_{-\infty}^\infty x^2 e^{-ax^2} dx = 1/(2a) \sqrt{\frac{\pi}{a}} \quad (1.9a)$$

$$\int_{-\infty}^\infty x^4 e^{-ax^2} dx = 3/(2^2 a^2) \sqrt{\frac{\pi}{a}} \quad (1.9b)$$

and so on, with repeated differentiation. If we now substitute (1.8) into (1.7) and use (1.5, 1.9) we find:

$$\begin{aligned} I_2 &= \sqrt{\frac{\pi}{a}} [1 - k^2/(4a) + k^4 3/(a^2 4! 2^2) - k^6 3.5/(a^3 6! 2^3) + \dots] \\ &= \sqrt{\frac{\pi}{a}} \left[ 1 - k^2/(4a) + (k^2/(4a))^2 1/2! - (k^2/(4a))^3 1/3! + \dots \right]. \end{aligned} \quad (1.10)$$

The term in the square brackets of (1.10) is just the series expansion for  $\exp[-k^2/(4a)]$ . (Here we have used the theorem that the order of integration and summation may be interchanged in expressions in which convergence is retained when each term is replaced by its absolute value.) Our result is:

$$\int_{-\infty}^{\infty} \cos kx e^{-ax^2} dx = \sqrt{\pi/a} e^{-k^2/4a}. \quad (1.11)$$

The right side of (1.11) is again a gaussian. Furthermore, the larger the wave number  $k$ , that is, the more rapidly  $\cos kx$  oscillates, the smaller the overlap integral (1.11) becomes. If  $a$  is large, the original gaussian is narrowly peaked and the overlap becomes broad in  $k$ .

Before concluding this section we need one more integral:

$$I_3 = \int_{-\infty}^{\infty} \sin kx e^{-ax^2} dx = 0. \quad (1.12)$$

Why is this integral zero? The reader will notice that whereas  $\cos kx = \cos(-kx)$ ,  $\sin kx = -\sin(-kx)$ ,  $\cos kx$  has *even symmetry*,  $\sin kx$  has *odd symmetry* about  $x = 0$ . Now, in (1.12) we are taking the overlap integral of an odd function ( $\sin kx$ ), with an even function  $\exp(-ax^2)$ . Such an overlap integral always vanishes because for every contribution coming from between  $x$  and  $x+dx$ , there is an equal but opposite contribution coming from between  $-(x+dx)$  and  $-x$ , the integrand is odd.

### 1.3 Complex Numbers

In the study of algebraic equations of quadratic and higher orders the reader will have encountered the use of  $i$ , the square root of  $-1$ . With  $i$  one constructs the two-fold set of complex numbers  $z$  out of real numbers. We write:

$$z = a + ib \quad (1.13a)$$

where  $a$  and  $b$  are real numbers.

Complex numbers can be represented as points in the “complex plane”, Fig. 1.1. The  $x$ -coordinate gives the real component of the number  $a$ , and the  $y$ -coordinate gives the imaginary component  $b$ . By “equality” of two complex numbers we mean that the real components of the two numbers are equal to each other, and the imaginary components are also equal:

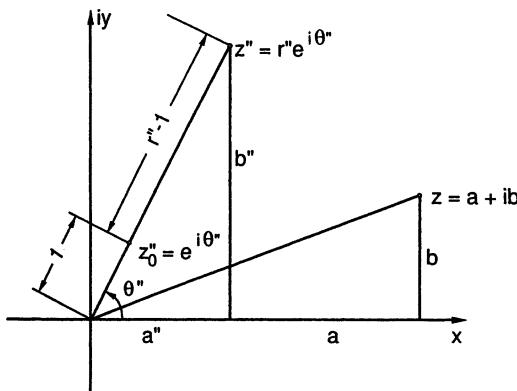
$$z = z' \quad \text{with} \quad (1.13b)$$

$$z = a + ib, \quad z' = a' + ib'$$

means

$$a = a' \quad b = b'.$$

In Fig. 1.1 we see that if  $z = z'$ ,  $z$  and  $z'$  are represented by the same point in the complex plane. For various algebraic manipulations with complex



**Fig. 1.1.** Two arbitrary complex numbers  $z, z''$  represented on the complex plane.  $r''$  is the absolute value of  $z''$ ,  $\theta''$  the phase of  $z''$ . Also shown is  $z_0$ , of absolute value unity,  $z_0'' = e^{i\theta''}$

numbers the reader is referred to books on intermediate algebra, and to the problems at the back of this chapter.

We generalize the well-known series representation for the exponential function of a real variable

$$e^x = 1 + x + x^2/2! + x^3/3! + \dots$$

to complex numbers

$$e^z = 1 + z + z^2/2! + z^3/3! + \dots, \quad (1.14)$$

where  $z$  is any complex number.

Let us see what happens when  $z$  is pure imaginary, that is,

$$z \equiv i\theta \quad (1.15)$$

where  $\theta$  is real. Remembering that  $i^2 = -1$ , we have from (1.14)

$$e^{i\theta} = 1 + i\theta - \theta^2/2! - i\theta^3/3! + \theta^4/4! \dots$$

Grouping the real and imaginary contributions together, we find

$$e^{i\theta} = 1 - \theta^2/2! + \theta^4/4! + \dots + i(\theta - \theta^3/3! + \dots).$$

We recall the series for the sine and cosine functions:  $\sin \theta = \theta - \theta^3/3! + \theta^5/5! \dots$  and  $\cos \theta = 1 - \theta^2/2! + \theta^4/4! \dots$ .

What we see is that

$e^{i\theta} = \cos \theta + i \sin \theta .$

(1.16)

In our geometric representation of complex numbers,  $\exp(i\theta'')$  would correspond to the point at the tip of a radius vector of length unity, and making an angle  $\theta''$  with the real axis. Then the  $x$  and  $y$  projections of the radius vector will just be the real and imaginary components of  $\exp(i\theta'')$ . We often write an arbitrary complex number  $z''$  as

$$\begin{aligned} z'' &= r'' e^{i\theta''} = a'' + ib'' \\ r'' &= \sqrt{a''^2 + b''^2}, \quad \tan \theta'' = b''/a''. \end{aligned} \quad (1.17)$$

$z''$  is to be represented by the end of a radius vector of length  $r''$ ,  $r''$  is referred to as the absolute value of  $z''$ ,  $r'' = |z''|$ ,  $\theta''$  as the phase. Figure 1.1 makes all these relationships clear.

The use of (1.17) for  $\exp(i\theta)$  in terms of trigonometric functions allows us to combine (1.11, 12):

$$\int_{-\infty}^{\infty} e^{ikx} e^{-ax^2} dx = \sqrt{\frac{\pi}{a}} e^{-k^2/4a}. \quad (1.18)$$

[Equation (1.18) can be derived directly by using Cauchy's Theorem, which lets us shift the contour of integration from the real axis to  $x - ik/(2a)$ . Complex variable theory leads to many new results in mathematics. Complex numbers also simplify many relationships used in physics. As we shall see in Chap. 6, complex variables are necessary in the formulation of quantum mechanics. Equation (1.18) is a result we shall use very often.]

## 1.4 The Dirac Delta Function

Let us first develop an intuitive understanding of the Dirac delta "function" through its simulation by sequences and then outline the rigorous treatment using the theory of generalized functions. We remember from the introduction that we are dealing with a spike-like function.  $\delta(x - a)$  has its spike at  $a$ . We can summarize the properties of  $\delta(x - a)$  as follows:

$$\int_{-\infty}^{\infty} F(x) \delta(x - a) dx = F(a) \quad (1.19a)$$

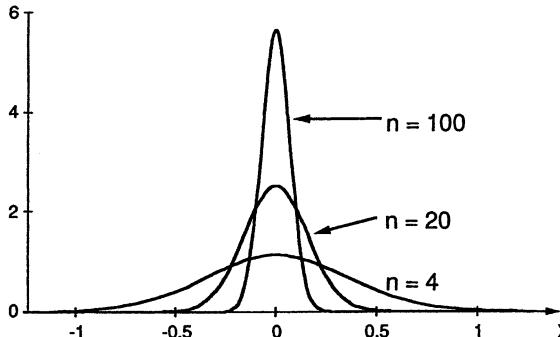
$$\int_{-\infty}^{\infty} \delta(x - a) dx = 1 \quad (1.19b)$$

$$\delta(x - a) = 0 \quad \text{if } x \neq a \quad (1.19c)$$

where  $F(x)$  is any good function.

Although  $\delta(x)$  is not a function in the ordinary sense, its action can be simulated by a sequence of functions. We obtain  $g_n(x), n = 0, 1, 2, 3 \dots$  by replacing the parameter  $a$  in (1.6) with the integer  $n$ :

$$g_n(x) = \sqrt{\frac{n}{\pi}} e^{-nx^2}. \quad (1.20)$$



**Fig. 1.2.** As  $n$  gets large the gaussian rapidly becomes a spike-like function. This function simulates the “Dirac delta function”. The area under all the curves is unity

We need to investigate the integral

$$\int_{-\infty}^{\infty} g_n(x)F(x)dx \quad \text{as } n \rightarrow \infty .$$

In Fig. 1.2 we show the behavior of  $g_n(x)$  as a function of  $x$  for increasing  $n$ . Clearly  $g_n(x)$  becomes higher and more localized. The area under the curve, however (the “strength” of the function), remains constant. Since with  $n$  large  $g_n(x)$  is a narrow spike at  $x = 0$ , all contributions to the integral must come from the immediate neighborhood of  $x = 0$ . We may think of  $F(x)$  as having the constant value  $F(0)$  in that neighborhood. Thus we write:

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} g_n(x)F(x)dx \cong \int_{-\infty}^{\infty} F(0)g_n(x)dx = F(0) \int_{-\infty}^{\infty} g_n(x)dx = F(0) ,$$

where in the last equality we have used (1.6).

Proceeding with the rigorous approach we define a sequence of good functions  $f_n(x)$  to be *regular* if for any good function  $F(x)$  the limit

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} f_n(x)F(x)dx$$

exists. Two regular sequences of good functions are *equivalent* if the limit is the same for each sequence. The sequences  $\exp(-x^2/4n)$  and  $\exp(-x^4/n^2)$ , for example, are equivalent. Finally, a *generalized function*  $f(x)$  is defined as a regular sequence  $f_n(x)$  of good functions, and two generalized functions are said to be equal if the corresponding sequences are equivalent. We therefore write for a generalized function  $f(x)$ :

$$\int_{-\infty}^{\infty} f(x)F(x)dx = \lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} f_n(x)F(x)dx . \quad (1.21)$$

Many ordinary functions,  $f_{\text{or}}(x)$ , can be viewed as generalized functions. All that is required is that at least one regular sequence of good functions  $f_n(x)$  exists for which

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} f_n(x) F(x) dx = \int_{-\infty}^{\infty} f_{\text{or}}(x) F(x) dx. \quad (1.22)$$

[It turns out that if  $f_{\text{or}}(x)$  is such that  $(1+x^2)^{-N} f_{\text{or}}(x)$  is absolutely integrable from  $-\infty$  to  $\infty$  for some  $N$ , then a sequence  $f_n(x)$  can be constructed.] Equation (1.22) greatly extends the scope of the theory of generalized functions.

We now prove that the sequences equivalent to  $g_n(x)$ , see (1.20), define a generalized function  $\delta(x)$  such that

$$\int_{-\infty}^{\infty} \delta(x) F(x) dx \equiv \lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} g_n(x) F(x) dx = F(0). \quad (1.23)$$

*Proof.* For any good function  $F(x)$

$$\begin{aligned} & \left| \int_{-\infty}^{\infty} (n/\pi)^{1/2} e^{-nx^2} F(x) dx - F(0) \right| = \left| \int_{-\infty}^{\infty} (n/\pi)^{1/2} e^{-nx^2} [F(x) - F(0)] dx \right| \\ & \leq |F'(x)|_{\max} \int_{-\infty}^{\infty} (n/\pi)^{1/2} e^{-nx^2} |x| dx \\ & = |F'(x)|_{\max} (n\pi)^{-1/2} \rightarrow 0 \quad \text{as } n \rightarrow \infty \end{aligned}$$

Here we have assumed that  $|F(x) - F(0)|$  is always less than the maximum value of the derivative in the range  $\infty > x > -\infty$ , times  $|x|$ , ( $|F'(x)|_{\max} |x|$ ).

## 1.5 Fourier Analysis

In Sect. 1.1 we indicated that Fourier analysis is concerned with decomposing functions into individual “waves”, that is, periodic components. In classical physics, that is, physics as it was formulated before the advent of quantum mechanics, physical phenomena were divided into particles and waves. Chapters 3–5 deal with waves as encountered in classical physics. Fourier analysis has always been a tool for working with these. Chapter 6 begins the exposition of quantum mechanics, where both particle and wave aspects can be attributed to the same phenomena. Here Fourier analysis is a key for understanding this duality.

For a system with boundaries at infinity we must consider an infinite number of possible wavelengths. We express a given good function  $F(x)$ ,  $\infty > x > -\infty$ , by an integral over all possible values of the wavevector  $k$ :

$$F(x) = 1/\sqrt{2\pi} \int_{-\infty}^{\infty} A(k) e^{ikx} dk . \quad (1.24)$$

We use the exponential form  $\exp(ikx)$  rather than decomposing into sines and cosines as a matter of convenience. If we allow the coefficients  $A(k)$  to also be complex, then there is no loss of generality, as  $A(k)$  then carries a phase.  $F(x)$  may be real or complex. These matters are examined in more detail in the problems at the end of this chapter. There we find that if, for instance,  $F(x)$  is a real and even function of  $x$ , the Fourier representation (1.24) may be written in terms of only  $\cos(kx)$  and real Fourier coefficients  $A(k)$ . We note that the absolute values  $|A(k)|$  are a measure of the relative importance of the various wave components  $(2\pi)^{-1/2} \exp(ikx)$  in the function  $F(x)$  being analyzed.

Equation (1.24) is useful only if some means exist of finding  $A(k)$  for any given function  $F(x)$ . When found, the function  $A(k)$  is called the Fourier transform of  $F(x)$ . There is one function for which we can find the Fourier transform with the knowledge we already have. Equation (1.18) will allow us to find the transform for the Dirac delta function. With this Fourier representation we shall be able to invert (1.24) and find the  $A(k)$  for any good function  $F(x)$ .

Replacing  $a$  by  $1/(4n)$ ,  $n$  again an integer, and interchanging  $k$  and  $x$  in (1.18) we find:

$$g_n(x) = \sqrt{n/\pi} e^{-nx^2} = 1/(2\pi) \int_{-\infty}^{\infty} e^{ikx} e^{-k^2/4n} dk . \quad (1.25)$$

We now write

$$g_n(x) = 1/\sqrt{2\pi} \int_{-\infty}^{\infty} b_n(k) e^{ikx} dk . \quad (1.26a)$$

Equation (1.26a) is in the same form as (1.24) with the desired transform  $b_n(k)$  given by

$$b_n(k) = 1/\sqrt{2\pi} e^{-k^2/4n} . \quad (1.26b)$$

Suppose now that any generalized function  $f(x)$  is given by a sequence  $f_n(x)$ , each member of the sequence having a Fourier representation

$$f_n(x) = 1/\sqrt{2\pi} \int_{-\infty}^{\infty} a_n(k) e^{ikx} dk . \quad (1.27a)$$

One can then show that the Fourier transforms of all equivalent sequences  $f_n(x)$ , namely  $a_n(k)$ , are also regular and equivalent sequences of good functions. Let  $a(k)$  be the generalized function equal to the sequences  $a_n(k)$ . That is, using (1.21)

$$\int_{-\infty}^{\infty} a(k)F(k)dk = \lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} a_n(k)F(k)dk , \quad (1.27b)$$

$a(k)$  will be taken as the definition of the Fourier transform of the generalized function  $f(x)$ .

Returning to the Fourier transform sequences of  $g_n(x)$ , namely the sequences  $b_n(k)$ , we have the obvious limit

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} e^{-k^2/4n} F(k)dk = \int_{-\infty}^{\infty} F(k)dk . \quad (1.28a)$$

We may, therefore, write

$$\int_{-\infty}^{\infty} (1/\sqrt{2\pi})F(k)dk = \lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} b_n(k)F(k)dk . \quad (1.28b)$$

It follows from (1.26a, 1.27, 1.28b) that

the FT of  $\delta(x)$  is  $b(k) = 1/\sqrt{2\pi}$   
 $\delta(x)$  and  $1/\sqrt{2\pi}$  are both generalized functions.

How can we understand this discovery that the Fourier transform of the delta function is a constant? If  $A(k)$  in (1.24) is a constant then at  $x = 0$  all the waves will be in phase and will pile up, tending to an infinite value. If  $x \neq 0$ , then while some waves ( $\cos kx + i \sin kx$ ) will have, say, positive real and imaginary contributions at some value of  $x$ , these will be cancelled by negative contributions at the same value of  $x$  coming from, say, waves  $(k \pm \pi/x)$ . The net result, it turns out, is that adding together all waves with amplitudes tending to be equal leads to a limit of zero, except at  $x = 0$ , where the sum blows up. In the excercises we ask the reader to show that the imaginary contribution to the integral defining  $\delta(x)$  vanishes by symmetry.

We can model the procedure for finding  $A(k)$  in general, i.e., inverting (1.24), on the procedure for projecting out of a given vector  $\mathbf{F}$  any one of the desired components. An arbitrary vector  $\mathbf{F}$  can be represented in a three dimensional cartesian system in terms of three mutually orthogonal unit vectors,  $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ , (replacing  $i, j, k$ ):

$$\mathbf{F} = A_1 \mathbf{v}_1 + A_2 \mathbf{v}_2 + A_3 \mathbf{v}_3 , \quad (1.29)$$

with

$$\mathbf{v}_i \cdot \mathbf{v}_j = \delta_{ij} \quad i, j = 1, 2, 3. \quad (1.30)$$

The symbol on the right side of (1.30) is called the Kronecker delta  $\delta_{ij}$ , and is unity when  $i = j$  and zero when  $i \neq j$ . Equation (1.30) thus expresses the *orthonormality* of the *basis vectors*  $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ . To find one of the components  $A_i$  in (1.29) we take the vector dot product of both sides of the equation with  $\mathbf{v}_i$ . This procedure projects out from  $\mathbf{F}$  the desired component. For example, taking the dot product with  $\mathbf{v}_1$  we get

$$\mathbf{v}_1 \cdot \mathbf{F} = A_1 \mathbf{v}_1 \cdot \mathbf{v}_1 + A_2 \mathbf{v}_1 \cdot \mathbf{v}_2 + A_3 \mathbf{v}_1 \cdot \mathbf{v}_3.$$

Using (1.30) we find

$$A_1 = \mathbf{v}_1 \cdot \mathbf{F}. \quad (1.31)$$

It is not so difficult to prove (Problem 1.19) that *any* good function  $F(x)$  can be expressed without error by a Fourier representation as in (1.24):

$$F(x) = 1/\sqrt{2\pi} \int_{-\infty}^{\infty} A(k) e^{ikx} dk. \quad (1.24)$$

We have suggested in Sect. 1.1 that this representation may be compared to the expansion of *any* vector  $\mathbf{F}$  in the basis  $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$  (1.29). The functions  $\exp(ikx)/\sqrt{2\pi}, \infty > k > -\infty$ , form a basis or complete set for good functions.

We now want to establish the orthonormality of this basis set of functions. To do so we rewrite (1.25), replacing  $x$  by  $(k - k')$  and  $k$  by  $x$ . This gives

$$1/2\pi \int_{-\infty}^{\infty} e^{i(k-k')x} e^{-x^2/4n} dx = g_n(k - k').$$

If we let

$$u_k(x) \equiv e^{ikx}/\sqrt{2\pi} \quad u_{k'}^*(x) \equiv e^{-ik'x}/\sqrt{2\pi} \quad (1.32a)$$

then:

$$\int u_{k'}^*(x) u_k(x) e^{-x^2/4n} dx = g_n(k - k') \quad (1.32b)$$

where recalling (1.23), one sees that  $g_n(k - k')$  is a sequence equal to  $\delta(k - k')$ .

Equation (1.32b) tends with  $n \rightarrow \infty$  to the orthonormality condition for  $u_k(x), u_{k'}(x)$ . As stated in Sect. 1.1 the condition involves taking the “overlap integral” between either one of the functions and the complex conjugate of the other. Equation (1.32b) is the analogue for exponential waves of the orthonormality of the basis vectors. A Dirac delta function  $\delta(k - k')$ , equal to the sequence  $g_n(k - k')$ , replaces the Kronecker deltas,  $\delta_{ij}$ , in (1.30). The similarity of the two conditions is seen by comparing the properties of the two kinds of deltas:

$$\int_{-\infty}^{\infty} \delta(k - k') dk' = 1 \quad \text{with} \quad \delta_{ij} = 1 \quad i = j$$

and

$$\delta(k - k') = 0 \quad \text{if} \quad k \neq k' \quad \text{with} \quad \delta_{ij} = 0 \quad i \neq j.$$

Again, if  $(k = k')$  the two waves in (1.32b) are everywhere in phase, and the overlap integral tends to  $\infty$  with  $n$ . If  $(k \neq k')$  there is a relative phase  $(k - k')x$ , varying with  $k'$ , leading to overall cancellation.

The factor  $\exp(-(x^2/4n)$  or a member of another equivalent sequence gives a prescription for carrying out the integration in (1.32b), and is, when  $n$  tends to infinity, a part of the definition of orthonormality for the exponential waves. Without this factor the integral would not be defined; the reader can easily check that no limit is approached when the range of integration is made larger and larger. As noted in Sect. 1.1 the need for mathematical definition has a physical counterpart. The factor  $\exp(-(x^2/4n)$  with  $n \rightarrow \infty$  simulates the fact that there is after all a beginning and an end to the wave components. The individual waves fade out as  $\exp(-(x^2/8n)$ . (In reality, they may either be absorbed, or one does not wait for their reflections to arrive.)

Following the projection procedure with vectors, we now multiply both sides of (1.24) by an element  $1/\sqrt{2\pi} \exp(-ik'x)$ , and by  $\exp(-(x^2/4n)$  and integrate over all  $x$ :

$$\begin{aligned} & 1/\sqrt{2\pi} \int_{-\infty}^{\infty} e^{-ik'x} e^{-x^2/4n} F(x) dx \\ &= 1/2\pi \int_{-\infty}^{\infty} e^{-ik'x} e^{-x^2/4n} \left[ \int_{-\infty}^{\infty} A(k) e^{ikx} dk \right] dx. \end{aligned}$$

We interchange the order of integration on the right (this procedure can be justified for good functions):

$$\begin{aligned} & 1/\sqrt{2\pi} \int_{-\infty}^{\infty} e^{-ik'x} e^{-x^2/4n} F(x) dx \\ &= 1/2\pi \int_{-\infty}^{\infty} A(k) \left[ \int_{-\infty}^{\infty} e^{i(k-k')x} e^{-x^2/4n} dx \right] dk. \end{aligned}$$

Further, interchanging sides, and simplifying with (1.32) we have:

$$\int_{-\infty}^{\infty} g_n(k - k') A(k) dk = 1/\sqrt{2\pi} \int_{-\infty}^{\infty} F(x) e^{-ik'x} e^{-x^2/4n} dx.$$

If we now let  $n \rightarrow \infty$  then we may use (1.23) to reduce the left side. Similarly, the right side may be reduced rather trivially by using (1.28a). The result is

if

$$F(x) = 1/\sqrt{2\pi} \int_{-\infty}^{\infty} A(k) e^{ikx} dk \quad A(k') = 1/\sqrt{2\pi} \int_{-\infty}^{\infty} F(x) e^{-ik'x} dx . \quad (1.24)$$

then

$$A(k') = 1/\sqrt{2\pi} \int_{-\infty}^{\infty} F(x) e^{-ik'x} dx . \quad (1.33)$$

Equation (1.33) is the desired Fourier transform of (1.24). We have put both equations side by side for comparison. Although the relationships look similar, indeed, except for a minus sign they are said to have a reciprocal relationship, it helps to initially interpret them as having quite different meanings. Equation (1.24) is the Fourier representation for the known function  $F(x)$ , that is, it gives  $F(x)$  as a sum over basis functions  $1/\sqrt{2\pi} \exp(ikx)$  with complex amplitudes  $A(k)$ . Equation (1.33), the Fourier transform, tells us how these amplitudes  $A(k)$  are to be found once  $F(x)$  is given. We note the similarity of (1.33) to (1.31) which projects out the desired vector component. Equation (1.33) projects out the desired basis function component.

As they stand, (1.24) and (1.33) have been shown to apply to good functions. They, therefore, also apply to the good sequences  $f_n(x), a_n(k)$  which are equal, respectively, to generalized functions  $f(x), a(k)$ . In summary:

if

$$f_n(x) = 1/\sqrt{2\pi} \int_{-\infty}^{\infty} a_n(k) e^{ikx} dk \quad a_n(k') = 1/\sqrt{2\pi} \int_{-\infty}^{\infty} f_n(x) e^{-ik'x} dx . \quad (1.27a)$$

then

$$(1.34)$$

and, we recall the definitions of  $f(x), a(k)$ :

$$\int_{-\infty}^{\infty} f(x) F(x) dx \cong \lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} f_n(x) F(x) dx \quad (1.21)$$

$$\int_{-\infty}^{\infty} a(k) F(k) dk \cong \lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} a_n(k) F(k) dk . \quad (1.27b)$$

As already stated, the sequences  $a_n(k)$  are equivalent if the sequences  $f_n(x)$  are, and the generalized function  $a(k)$  is, by definition, the FT of the generalized function  $f(x)$ . We have seen in connection with (1.22) that there exists a wide class of functions that can be viewed as generalized functions. The combination of good functions and all generalized functions includes all the functions for which a theoretical physicist is likely to need Fourier representations.

The reader will wonder whether the evaluation of the functions  $A(k)$  by means of (1.33) or (1.34) is always an easy matter. In practice the situation is not that difficult. Apart from the Dirac delta function there are two functions which, perhaps, one is called upon to Fourier analyze more frequently than any other. These are the gaussian and the “square wave”. We already have in (1.18, 25) the representation of a gaussian. The FT is again a gaussian – a well-known result. The “square wave” is analyzed in Examples 1.1 and 1.2. Some other simple transforms will be found in the Examples and Problems. Many transforms can be found in published tables. Some of these have been worked out by using complex variable theory<sup>1</sup>. A different approach is to use numerical methods, usually with the help of a computer. The reader will find an illustration in Example 1.2, and a computing problem as a suggested laboratory exercise at the end of Problems.

## 1.6 Hilbert Space

In Sect. 1.1 we introduced the concept of infinite function space, which while encompassing infinite dimensions is built up in the same manner as the familiar three dimensional vector space. As we stated, the ordinary unit vectors  $\mathbf{i}, \mathbf{j}, \mathbf{k}$  form a complete basis. This means that in ordinary vector space *any* vector can be described as a linear combination of these three basis vectors, with suitable coefficients fixing the size of the corresponding components. In the same way, a set of basis functions, such as  $u_k(x) = 1/\sqrt{2\pi} \exp(ikx)$ ,  $\infty > k > -\infty$ , can be considered complete if the functions of physical interest can be linearly expressed in terms of this basis. Many other examples of basis built from orthogonal functions of  $x$  as arise in the theory of differential equations find application in physics. The question of completeness was studied by Hilbert in a mathematically rigorous fashion, exceeding our needs. The interested reader is referred to the book by Byron and Fuller listed at the end of the chapter. For the exponential functions  $u_k(x)$  of this chapter we may be satisfied with the proof of their completeness outlined in Problem 1.19. Infinite function spaces in general that are complete are

---

<sup>1</sup> The FTs of many functions which are ordinary functions that can be viewed as generalized functions can be obtained directly from (1.33) for good functions. Sometimes, however, one must find a suitable sequence satisfying (1.22), see Example 1.2.

called *Hilbert spaces*. Hilbert space is the setting for quantum mechanics, and the term, meaning the space of an infinite complete set of basis functions is prevalent in theoretical physics.

It will be helpful by way of a summary to point out explicitly the analogues between  $N$ -dimensional vector space and Hilbert space. In Problem 1.20 we give a simple development, showing how the continuum expressions governing Hilbert space arise naturally from their analogues in  $N$ -dimensional vector space. The latter, the reader will perceive is an obvious generalization of three dimensional space.

<i>N</i> -Dimensional Vector Space	Hilbert Space
The <i>basis</i> of the space are the unit vectors: $\mathbf{v}_i; \quad i = 1, 2, 3, \dots, N$ with $N = 3$ $\mathbf{v}_1 = \mathbf{i}, \mathbf{v}_2 = \mathbf{j}, \mathbf{v}_3 = \mathbf{k}$ .	The <i>basis</i> of the space are the functions: $u_k(x), \quad \infty > k > -\infty$ Discretely numbered functions $u_n(x), \quad n = 1, 2, 3, \dots$ are also possible, see Chap. 6.
The <i>orthonormality</i> condition is: $\mathbf{v}_i \cdot \mathbf{v}_j = \delta_{ij} \quad i, j = 1, 2, \dots, N$ On the left is the vector dot product, on the right the Kronecker delta.	The <i>orthonormality</i> condition is: $\int_{-\infty}^{\infty} u_k^*(x) u_{k'}(x) dx = \delta(k - k')$ On the left is the overlap integral, on the right the Dirac delta function.
<i>Completeness</i> means that any vector $\mathbf{F}$ in the space can be written as: $\mathbf{F} = \sum_{i=1}^N A_i \mathbf{v}_i$	<i>Completeness</i> means that any function $F(x)$ in the space can be written as: $F(x) = \int_{-\infty}^{\infty} A(k) u_k(x) dk$ or, if the basis is discrete, as: $F(x) = \sum_n A_n u_n(x)$ where the $A(k), A_n$ are the amplitudes of the basis functions $u_k(x), u_n(x)$ .
<i>Projection Formula</i> $A_i = \mathbf{v}_i \cdot \mathbf{F}$ with the right side being a vector dot product.	<i>Projection Formula</i> $A(k) = \int_{-\infty}^{\infty} u_k(x) F(x) dx$ with the right side being an overlap integral.

**Example 1.1.** Find the FT of the square wave function  $F(x)$  defined by

$$\begin{aligned} F(x) &= 0 & x \geq a, x \leq -a \\ F(x) &= v & a > x > -a \end{aligned}$$

and discuss the result.

*Solution.* We expand the exponential in (1.24), the Fourier representation for good functions. [We postpone discussion of the fact that  $F(x)$  is not a good function, but rather an “ordinary function that may be viewed as a generalized function” until the end of Example 1.2.] We find:

$$A(k) = 1/\sqrt{2\pi} \int_{-\infty}^{\infty} (\cos kx - i \sin kx) F(x) dx .$$

Again, because  $F(x)$  is even in  $x$  while  $\sin kx$  is odd the integral  $\int_{-\infty}^{\infty} \sin(kx) \times F(x) dx$  vanishes.

Making direct use of the definition of  $F(x)$ , we can now write

$$\begin{aligned} A(k) &= 1/\sqrt{2\pi} \int_{-\infty}^{\infty} F(x) \cos kx dx \\ &= v/\sqrt{2\pi} \int_{-a}^a \cos kx dx \\ &= v/\sqrt{2\pi} [\sin ka - \sin(-ka)]/k \\ &= \sqrt{2/\pi}(v/k) \sin(ka) . \end{aligned}$$

We are asked to discuss this result: We first notice that  $A(k)$  is even in  $k$ . This allows us to collect together the contribution of  $+k$  and  $-k$  for every  $k$  in the original Fourier representation, (1.24), and to write a new representation valid for all even functions:

$$F(x) = \sqrt{2/\pi} \int_0^{\infty} A(k) \cos kx dk ,$$

where we have used  $\exp(ikx) + \exp(-ikx) = 2 \cos kx$ .

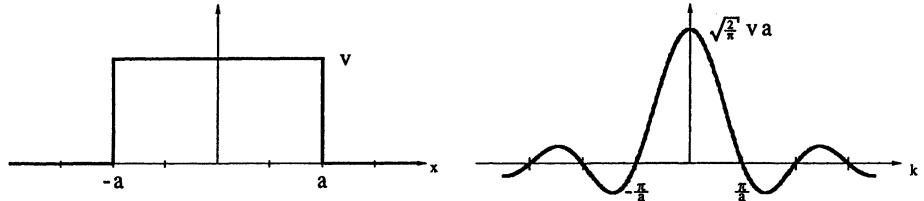
Turning now to the result for  $A(k)$ , if we ask ourselves at what values of  $k$  we expect the coefficients  $A(k)$  to be most pronounced, we might guess this to be in the vicinity of  $k = \pi/2a$ , because the central lobe of  $\cos \pi x/2a$  most nearly resembles the given function, and can be expected to have the greatest overlap integral with it. In general the long waves ( $ka \ll 1$ ) will assure correct long range representation of a function, and short waves will take care of any rapid changes. Recalling that

$$\lim_{k \rightarrow 0} \sin ka/k = a ,$$

we see that in the present case (Fig. 1.3), we get the largest contributions to the transform over the range

$$\pi/a \stackrel{\sim}{>} |k| > 0 .$$

For large  $|k|$ ,  $A(k)$  goes to zero, but rather slowly. The very short waves are needed to help produce the infinitely sharp edges at  $x = \pm a$  (as we verify in the next example).



**Fig. 1.3.** The square wave and its FT, see Example 1.1

Lastly, we notice that  $\Delta k \approx 1/a$  is the range in  $k$  over which we get the most important Fourier components, and  $\Delta x \approx a$  is the range of the original function, so that

$$\boxed{\Delta x \Delta k \approx 1}.$$

This is a most important result, which can be extended to any function and its Fourier transform. A large spread in ordinary space leads to a narrow spread in “ $k$ -space” and vice versa. Later we shall establish this result more precisely in connection with the gaussian function.

**Example 1.2.** Use the computer to verify that, indeed, a Fourier cosine representation with  $\sqrt{2/\pi}(\sin k)/k$  yields a square wave extending from  $x = -1$  to  $x = +1$  with unit height. Show explicitly the contribution of the short waves.

*Solution.* According to Example 1.1

$$\begin{aligned} F(x) &= \sqrt{2/\pi} \int_0^\infty A(k) \cos kx dk \\ &= 2/\pi \int_0^\infty (\sin k)/k \cos kx dk \cong I_1 + I_2 \end{aligned}$$

where we will take  $I_1$  to be the main contribution ending at  $k = 3\pi/4$ , and  $I_2$  the short wave contribution ending at a large number, say  $32\pi$ . We have

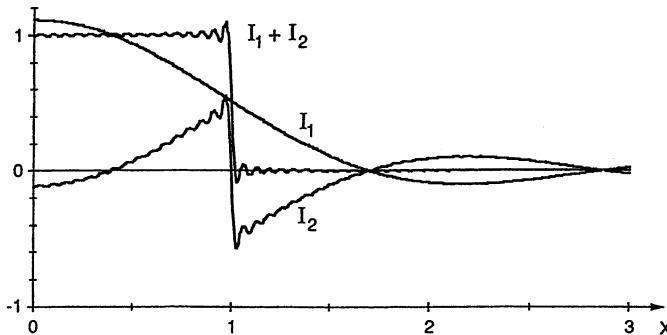
$$\begin{aligned} I_1 &\equiv 2/\pi \int_0^{3\pi/4} (\sin k)/k \cos kx dk \\ I_2 &\equiv 2/\pi \int_{3\pi/4}^{32\pi} (\sin k)/k \cos kx dk. \end{aligned}$$

Numerically, we take  $\delta k = 0.05$ , and notice that  $(3\pi/4)/0.05 = 47.124$ . We then evaluate  $I_1, I_2$  at the points  $x = 0.01m$ ,  $m = 0, 1, 2, \dots, 300$ , with limits on  $n$  as shown in the summations:

$$I_1 = (2 \times 0.05)/\pi \left[ 1/2 + \sum_{n=1}^{47} (\sin 0.05n)/(0.05n) \cos(0.05nx) \right]$$

$$I_2 = (2 \times 0.05)/\pi \left[ \sum_{n=48}^{2010} (\sin 0.05n)/(0.05n) \cos(0.05nx) \right].$$

Figure 1.4 shows  $I_1, I_2$ , and  $(I_1 + I_2)$  as functions of  $x$  for  $x > 0$ . We notice that  $I_1$  indeed represents the smoothed out square wave. The sharp features arise from the short wave contribution  $I_2$ .



**Fig. 1.4.** Fourier synthesis of a square wave function for  $x > 0$ . The principal contribution,  $I_1$ , comes from  $3\pi/4 > k > 0$ .  $I_2$  for the short wavelengths, viz.  $32\pi > k > 3\pi/4$ , and  $(I_1 + I_2)$  are also shown. The rapid oscillations are associated with Gibbs' phenomenon, see Example 1.2

The sudden drop at  $x = 1$  in  $F(x)$  give rise to the rapid oscillations in the Fourier representation at this point. The characteristic wavelength  $\lambda$  of these oscillations is  $\lambda \sim 2\pi/k_0$  where  $k_0$  is the largest  $k$  (cut-off) in our numerical calculation. If we could increase  $k_0$  to infinity the oscillations would not entirely disappear, but degenerate into small but finite over-shoots of zero width to the values of the function  $F = 1$  and  $F = 0$  at  $x = 1$  (and at  $x = -1$ ). These over-shoots will occur whenever the represented function has infinite slopes and are referred to as Gibbs' phenomenon.

We note several points about the phenomenon. Because the over-shoots are of *finite* height and zero width they will have no effect if the represented function is a factor in an integral.

Second, the phenomenon is avoided altogether if  $F(x)$  is viewed as a generalized function, i.e., taken to be “equal” to a sequence  $f_n(x)$ , where in  $f_n(x)$  the discontinuities at  $x = \pm 1$  are “smudged out” over a width of order  $1/n$ . All that is then required is that in Fourier representing  $f_n$ , the cut-off wave vector  $k_0 \gg n$ .

Lastly, the analytic form obtained in Example 1.1 for  $A(k)$  will give the same end result for test integrals defining the generalized transform (1.27b) as the procedure of finding  $a_n(k)$  and letting  $n \rightarrow \infty$ . Thus, the FT of “an ordinary function that can be viewed as a generalized function” rather than being a good function itself can often be obtained directly from (1.33) (see footnote to the last section), although here some small difficulty did arise. Direct results, such as the analytic form for  $A(k)$  here, or the results given in Problems 1.11 – 13c, are generally used in physical calculations. However, if, for example, derivatives of discontinuous functions (polynomials in the transform functions) are involved, we do need the full theory of generalized functions.

**Example 1.3.** The delta function has an orthonormal property analogous to that in (1.32b) of the functions  $u_k(x)$ , namely,

$$\int_{-\infty}^{\infty} \delta(x-a)\delta(x-b)dx = \delta(a-b) .$$

Use the theory of generalized functions to give a formal proof of this intuitive result. (In the context of quantum mechanics (Chap. 6), it is this property rather than (1.19b) which implies “normalization” of the delta function.)

*Solution.* Equations (1.21) and (1.32b) indicate that for the left side of the relation we may consider the limit as  $l \rightarrow \infty$ , of  $I_{lmn}$ , where

$$I_{lmn} = \int_{-\infty}^{\infty} e^{-x^2/4l} g_m(x-a)g_n(x-b)dx .$$

Fourier representing  $g_m(x-a)$ ,  $g_n(x-b)$  as in (1.26b) and rearranging the order of integrations, one has

$$\begin{aligned} (2\pi)^2 I_{lmn} &= \\ &\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[ e^{-x^2/4l} e^{ik_1(x-a)} e^{ik_2(x-b)} \right] dx e^{-k_1^2/4m} e^{-k_2^2/4n} dk_1 dk_2 \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} 2\pi g_l(k_1 + k_2) e^{-ik_1 a} e^{-ik_2 b} e^{-k_1^2/4m} e^{-k_2^2/4n} dk_1 dk_2 . \end{aligned}$$

Proceeding to the limit  $l \rightarrow \infty$  ( $\lim l \rightarrow \infty, I_{lmn} \equiv I_{mn}$ ) one gets from (1.23) that

$$2\pi I_{mn} = \int_{-\infty}^{\infty} e^{ik_2(a-b)} e^{-k_2^2(n+m)/4nm} dk_2 ,$$

or  $I_{mn} = g_{nm/n+m}(a - b)$ , which is indeed a sequence for  $\delta(a - b)$ , as  $n$ , say, gets large compared to  $m$ , and then  $m$  also gets large, or as  $m$  and  $n$  get large together.

**Example 1.4.** In Chap. 6, in connection with the second postulate of quantum mechanics, it is stated that the solution to the equation

$$xu_\xi(x) = \xi u_\xi(x) \quad (6.20)$$

is

$$u_\xi(x) = \delta(x - \xi) \quad (6.21)$$

that is, the Dirac delta function centered at  $\xi$ . Formulate this statement in terms of the theory of generalized functions, and give a proof.

*Solution.* When  $u_\xi(x)$  and  $xu_\xi(x)$  are understood to be generalized functions, (6.20) means that:

$$\int_{-\infty}^{\infty} F(x)xu_\xi(x)dx = \int_{-\infty}^{\infty} F(x)\xi u_\xi(x)dx \quad (6.20a)$$

for any good function  $F(x)$ . Similarly (6.21) means that

$$\int_{-\infty}^{\infty} F(x)u_\xi(x)dx = \int_{-\infty}^{\infty} \delta(x - \xi)F(x)dx, \quad (6.21a)$$

again with generalized functions appearing in the integrands along with  $F(x)$ .

As proof of (6.20a) we have, replacing  $F(x)$  by  $xF(x)$  in (6.21a), also a good function:

$$\begin{aligned} \int_{-\infty}^{\infty} xF(x)u_\xi(x)dx &= \int_{-\infty}^{\infty} xF(x)\delta(x - \xi)dx \\ &= \lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} xF(x)g_n(x - \xi)dx \\ &= \lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} \xi F(x)g_n(x - \xi)dx \\ &= \int_{-\infty}^{\infty} F(x)\xi u_\xi(x)dx, \end{aligned}$$

where the second and third line both have  $\xi F(\xi)$  as limit, and the last equality follows from (6.21a).

**Example 1.5.** Show that the FT of  $xf(x)$  is  $\text{i}a'(k)$  when  $a(k)$  is the FT of  $f(x)$  and, hence, that the FT of  $x$  is  $\sqrt{2\pi}\text{i}\delta'(k)$ , all being regarded as generalized functions of  $x$  or  $k$ , respectively. The primes indicate derivatives with respect to the indicated variable.

*Solution.* We remember that the FT  $a_n(k)$  of a regular sequence  $f_n(x)$  defines  $a(k)$ , the FT of  $f(x)$ . Differentiating both sides of (1.34) with respect to  $k$ , we find

$$a'_n(k) = 1/\sqrt{2\pi} \int_{-\infty}^{\infty} -\text{i}x f_n(x) e^{-\text{i}kx} dx .$$

Inverting, we have

$$-\text{i}x f_n(x) = 1/\sqrt{2\pi} \int_{-\infty}^{\infty} a'_n(k) e^{\text{i}kx} dk ,$$

showing that the FT of  $xf_n(x)$  is  $\text{i}a'(k)$ , in the limit  $n \rightarrow \infty$ . Further, if we choose  $f_n(x)$  to be  $\exp(-x^2/4n)$ , one of the sequences equivalent to unity, then by (1.18):

$$a_n(k) = 1/\sqrt{2\pi} \int_{-\infty}^{\infty} e^{-x^2/4n} e^{-\text{i}kx} dx = \sqrt{2n} e^{-nk^2} ,$$

or  $a_n(k)$  is a sequence of  $\sqrt{2\pi}\delta(k)$  and  $a'_n(k)$  a sequence giving  $\sqrt{2\pi}\delta'(k)$ . It follows that the FT of  $x$  is  $\text{i}\sqrt{2\pi}\delta'(k)$  from the first part of the example. (The formulation of the theory of generalized functions in terms of good functions allows the extension of these results to arbitrary powers of  $x$  and arbitrarily high derivatives of the  $\delta$  function.)

## Problems

**1.1** Find a general formula for

$$1/(2n)! \int_{-\infty}^{\infty} x^{2n} e^{-ax^2} dx .$$

**1.2** A fairly good function is one that is everywhere differentiable any number of times, and such that it and all its derivatives are  $O(|x|^N)$  as  $x \rightarrow \infty$  for some  $N$ . Show that the product of a good function and a fairly good function is a good function.

**1.3** Write  $1/(a + ib)$  as the sum of a real and imaginary number.

**1.4** a) Multiply  $z_1 = A_1 e^{i\theta_1}$  by  $z_2 = A_2 e^{i\theta_2}$  ( $A_1, A_2, \theta_1, \theta_2$  real). Represent the whole operation on the complex plane.

b) Consider the special case  $A_1 = A_2, \theta_1 = -\theta_2$ . What can you say about the relationship between  $z_1$  and  $z_2$ ?

**1.5** a) Prove de Moivre's Theorem:

$$(\cos \theta + i \sin \theta)^n = \cos n\theta + i \sin n\theta .$$

b) Find formulas for  $\cos 3\theta$  and  $\sin 3\theta$ .

**1.6** a) Show that if the derivative  $f'(x)$  of a generalized function  $f(x)$  is defined by the sequence  $f'_n(x)$ , where  $f(x)$  is defined by  $f_n(x)$ , then

$$\int_{-\infty}^{\infty} f'(x) F(x) dx = - \int_{-\infty}^{\infty} f(x) F'(x) dx .$$

b) Find  $\int_{-\infty}^{\infty} x e^{-x^2/n} \delta'(x) dx .$

c) Sketch  $g'_n(x)$  for large  $n$ . What does it do as part of an integrand?

**1.7** In the representation for the  $\delta$  function sequence  $g_n$ , (1.26a), there appear complex quantities on the right, whereas the left side is real. Can you explain this paradox?

**1.8** An incoming FM radio signal has an amplitude given by  $E(t)$ , with  $t$  the time. Write  $E(t)$  as the superposition of all possible angular frequencies  $\omega$ . Make a rough plot of how you suppose the absolute value of the Fourier coefficients  $|A(f = \omega/2\pi)|$  looks as a function of the frequency  $f$ .

Put numbers on the frequency axis.

*Hint:* Look at your FM dial, and recall that the side-bands of any one channel extend through about  $\pm 75$  kHz.

**1.9** In the text it is stated in connection with the Fourier representation (1.24), "We use the exponential form  $\exp(ikx)$  rather than decomposing into sines and cosines as a matter of convenience. If we allow the amplitudes  $A(k)$  to also be complex functions, then there is no loss of generality, as  $A(k)$  then carries a phase.  $F(x)$  may be real or complex."

a) Find a general relationship between  $A(k)$  and  $A(-k)$  so that  $F(x)$  will be real.

b) Write  $A(k) = |A_k| \exp(i\theta_k)$ , where  $|A_k|$  is the amplitude and  $\theta_k$  is the phase, and find the

Fourier representation for  $F(x)$  in terms of these quantities, assuming  $F(x)$  real. Now, convert your representation to sines and cosines and show that, indeed, enough parameters are provided with no loss of generality. (The functions  $\sin kx, \cos kx$  for  $\infty > k > 0$  form a complete basis, and for each harmonic function and value of  $k$  one parameter is needed. It may be helpful to discretize the given range of  $k$  into small but finite intervals.)

c) Show that when  $F(x)$  is complex, again enough parameters are provided for no loss of generality.

**1.10** a) Show that any function can be written as the sum of an even function,  $f_e(x)$  and an odd function  $f_o(x)$ .

b) Show that we may write the Fourier representation (1.24) as

$$F(x) = 1/\sqrt{\pi} \int_0^\infty C(k) \cos kx dk + 1/\sqrt{\pi} \int_0^\infty S(k) \sin kx dk .$$

Find  $C(k)$  and  $S(k)$  in terms of  $A(k)$ .

c) Find Fourier expansions for  $f_e(x)$  and  $f_o(x)$  in part (a) in terms of  $A(k), A(-k)$  and then in terms of  $C(k)$  and  $S(k)$ .

d) Find  $\int_{-\infty}^{\infty} \cos k'x \cos kx e^{-x^2/4n} dx$  and, then, a simple formula for  $C(k)$  in terms of  $f_e(x)$ .

e) Find  $\int_{-\infty}^{\infty} \sin k'x \cos kx e^{-x^2/4n} dx$ .

**1.11** a) Find  $A(k)$ , the FT of the function

$$\begin{aligned} F(x) &= e^{-ax} & x \geq 0 \\ &= 0 & x < 0 . \end{aligned}$$

b) Represent  $F(x)$  as a Fourier integral over sines and cosines.

**1.12** a) Find a set of tables of the Fourier transform and determine the transform of  $(\sin nx)/\pi x$ .

b) Refer to Example 1.1, and find a way of corroborating your findings.

c) Show that

$$\lim_{n \rightarrow \infty} (\sin nx)/\pi x = \delta(x) .$$

Discuss.

d) Suggest two equivalent sequences for  $\delta(x)$ .

**1.13** a) Indicate by using the techniques of Problem 1.12, that is, by looking at the FT sequences, that the sequence

$$f_n(x) \equiv 1/(n\pi) [x^2 + n^{-2}]^{-1}$$

is equal to  $\delta(x)$ .

b) Show that

$$\begin{aligned} & \lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} F(x)/(x + i/n) dx \\ &= \lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} F(x)x/(x^2 + n^{-2}) dx - i\pi \int_{-\infty}^{\infty} F(x)\delta(x) dx . \end{aligned}$$

c) If  $P \int_{n-\infty}^{\infty} [F(x)/x] dx \equiv \lim_{n \rightarrow \infty} \int_{-\infty}^{-1/n} [F(x)/x] dx + \int_{+1/n}^{+\infty} [F(x)/x] dx$

show that

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} F(x)x/(x^2 + n^{-2}) dx = P \int_{-\infty}^{\infty} [F(x)/x] dx .$$

d) Show that the FT of the sequence of *good* functions

$$f_n(x) = (1/n\pi)e^{-x^2/n^2} [x^2 + n^{-2}]^{-1}$$

is given by

$$a_n(k) = 1/4\sqrt{2/\pi} \left[ e^{-k/n} \operatorname{Erfc}(n^{-2} - nk/2) + e^{k/n} \operatorname{Erfc}(n^{-2} + nk/2) \right] e^{-n^{-4}}$$

where  $\operatorname{Erfc}(y) = 2/\sqrt{\pi} \int_y^{\infty} e^{-x^2} dx$  for  $\infty > y > -\infty$ . Why do we say “indicate” in (1.13a)?

**1.14** (Parseval’s theorem for good functions.) Show that if  $F_1(x), F_2(x)$  are good functions and  $A_1(k), A_2(k)$  their FTs, then

$$\int_{-\infty}^{\infty} A_1(k)A_2(k) dk = \int_{-\infty}^{\infty} F_1(-x)F_2(x) dx .$$

**1.15** a) Show that

$$\int_{-\infty}^{\infty} F(x)H(y-x) dx = \int_{-\infty}^{\infty} A(k)B(k)e^{iky} dk ,$$

where  $A(k), B(k)$  are the FTs of  $F(x)$  and  $H(x)$ , respectively. This result is known as the “Faltungstheorem”.

b) Solve the integral equation

$$G(y) = \int_{-\infty}^{\infty} F(x)H(y-x) dx$$

where  $F(x)$  is the unknown function,  $G(x), H(x)$  are known.

**1.16** The correlation function for a random process is defined for a random variable  $x$  by

$$C(\tau) = \lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} x(t)x(t - \tau)e^{-t^2/4n} dt .$$

By representing  $x(t)$  as a Fourier integral

$$x(t) = \sqrt{2\pi} \int_{-\infty}^{\infty} A(f)e^{2\pi ift} df$$

show that with  $x(t)$  real

$$C(\tau) = 2\sqrt{2\pi} \int_{-\infty}^{\infty} |A(f)|^2 \cos(2\pi f\tau) df .$$

**1.17** a)<sup>2</sup> Show that the diffusion equation

$$\partial \varrho(t, x) / \partial t = D \partial^2 \varrho(t, x) / \partial x^2$$

has solutions of the form (that stay finite as  $x \rightarrow \pm\infty$ )

$$\varrho_k = e^{(ikx - k^2 Dt)} .$$

b) In general:

$$\varrho = 1/\sqrt{2\pi} \int_{-\infty}^{\infty} A(k)e^{(ikx - k^2 Dt)} dk .$$

Find  $A(k)$  if at  $t = 0$

$$\varrho(0, x) = \delta(x - x_0) .$$

c) Show that for the initial condition given in (b), the result obtained gives

$$\varrho(x, t) = 1/\sqrt{4\pi Dt} e^{-(x-x_0)^2/4Dt} .$$

**1.18** A member  $f_n(x)$  of the sequence corresponding to a suitable ordinary function  $f_{\text{or}}(x)$ , (1.22), can be constructed as

$$f_n(x) = \int_{-\infty}^{\infty} f_{\text{or}}(t) S\{n(t - x)\} n e^{-t^2/n^2} dt$$

---

<sup>2</sup> Problem 1.17 is important for Chap. 6 on wave mechanics.

where, what Lighthill calls the “smudge function”,  $S(y)$  is given by

$$S(y) = e^{-1/(1-y^2)} \left[ \int_{-1}^1 e^{-1/(1-z^2)} dz \right]^{-1} \quad 1 > y > -1$$

$$S(y) = 0 \quad y < -1, y > 1.$$

- a) How is  $S(y)$  normalized?
- b) Over what interval is  $f_{\text{or}}(t)$  averaged (“smudged”) out?
- c) What is the purpose of the factor  $e^{-t^2/n^2}$ ?

**1.19 Completeness.** The functions  $\exp(ikx)/\sqrt{2\pi}, \infty > k > -\infty$ , form a complete set if the Fourier representation

$$1/\sqrt{2\pi} \int_{-\infty}^{\infty} A(k) e^{ikx} dk \quad \text{where} \quad A(k) = 1/\sqrt{2\pi} \int_{-\infty}^{\infty} F(x) e^{-ikx} dx$$

tends to  $F(x)$  at all  $x$ , for any good function  $F(x)$ .

To verify completeness, substitute the given expression for  $A(k)$  into the given representation, and verify that this yields  $F(x)$  at all  $x$ .

**1.20 Overlap Integrals.** Two arbitrary vectors  $\mathbf{F}_A, \mathbf{F}_B$  can be represented in  $N$  dimensional space, generalizing (1.29), as:

$$\mathbf{F}_A = \sum_{i=1}^N A_i \mathbf{v}_i, \quad \mathbf{F}_B = \sum_{i=1}^N B_i \mathbf{v}_i \quad i = 1, 2, 3, \dots N$$

where  $\mathbf{v}_i$  is an orthonormal vector basis as in (1.30). The vector dot product becomes:

$$\mathbf{F}_A \cdot \mathbf{F}_B = \sum_{i=1}^N A_i B_i,$$

the sum over products of corresponding components.

We wish, now, to extend these concepts to “function space”, working with a discrete formulation that can then be easily taken to its continuum limit. Thus, we choose some finite increment  $\Delta$  in  $x$ , and consider an allowed range for  $x$ :

$$\frac{-N\Delta}{2} < x \leq \frac{N\Delta}{2}$$

where  $N$  is a large even number. We begin by working with  $N$  “square wave” functions,  $h_i(x)$ :

$$h_i(x) = 1/\sqrt{\Delta} \quad x_{i-1} < x \leq x_i$$

$$= 0 \quad \frac{N\Delta}{2} \geq x > x_i, \quad x_{i-1} > x > \frac{-N\Delta}{2}$$

where  $x_i = (i)\Delta$  and  $i = \frac{-N}{2} + 1, \frac{-N+2}{2}, \dots \frac{N}{2}$ .

The  $h_i(x)$  form a basis; in the given interval any function of  $x$ ,  $F_A(x)$ ,  $F_B(x)$ ... can be represented in terms of these functions with increasing precision as  $N \rightarrow \infty$ ,  $\Delta \rightarrow 0$ :

$$F_A(x) \cong \sum_{i=-N/2+1}^{N/2} A_i h_i(x); \quad F_B(x) \cong \sum_{i=-N/2+1}^{N/2} B_i h_i(x).$$

Moreover, the functions  $h_i(x)$  are independent (can not be written as linear combinations of each other). Indeed, they exist only at their own individual locations; they do not overlap in space. This last property can be best described analytically by defining an *overlap integral* for two functions, which is going to vanish if the functions selected are  $h_i(x)$ ,  $h_j(x)$ ,  $i \neq j$ . We have:

$$I_{ij} = \int_{-N\Delta/2}^{N\Delta/2} h_i(x) h_j(x) dx = \delta_{ij}$$

where the integral over  $x$  defines the overlap integral  $I_{ij}$  of any two functions  $h_i(x)$ ,  $h_j(x)$  whatever, while the Kronecker delta,  $\delta_{ij}$ , on the right expresses the orthonormality of the “square wave functions” just defined.

i) We can now anticipate that the overlap integral of the two arbitrary functions,  $F_A(x)$ ,  $F_B(x)$  will have the same form as the dot product of the vectors  $\mathbf{F}_A$  and  $\mathbf{F}_B$  above. Show by expanding  $F_A(x)$ ,  $F_B(x)$  in square wave functions  $h_i(x)$  that

$$\int_{-N\Delta/2}^{N\Delta/2} F_A(x) F_B(x) dx = \sum_{i=-N/2+1}^{N/2} A_i B_i$$

i.e. the sum of products of corresponding components (expansion coefficients).

ii) Show that

$$A_i \cong \int_{-N\Delta/2}^{N\Delta/2} F_A(x) h_i(x) dx,$$

the overlap integral of  $F_A(x)$  with the basis member  $h_i(x)$ , in analogy with the projection of an arbitrary vector on a unit vector. Does  $A_i \cong \sqrt{\Delta} F_A(x_i)$ ?

iii) Suppose that we can find a new orthonormal basis, that is, with the new overlap integral  $I_{ij} = \delta_{ij}$  again. The new basis.  $u_i(x)$ , could, for example, be comprised of the  $N$  functions

$$\sqrt{\frac{2}{N\Delta}} \sin k_j x, \sqrt{\frac{2}{N\Delta}} \cos k_j x, \quad \text{with} \quad k_j = \frac{2\pi j}{N\Delta}, j = 1, 2, 3, \dots, \frac{N}{2}.$$

Show, this time be expanding in  $u_i(x)$ , that

$$\int_{-N\Delta/2}^{+N\Delta'/2} F_A(x)F_B(x)dx = \sum_{i=-N/2+1}^{N/2} \widetilde{A}_i \widetilde{B}_i$$

with  $\widetilde{A}_i$ ,  $\widetilde{B}_i$  the new expansion coefficients.

Since, for the functions there is only one overlap integral (“dot product”)  $\int F_A(x)F_B(x)dx$ , the sum  $\sum A_i B_i$  is independent of the choice of basis. May we characterize the linear transformation expressing functions  $u_i(x)$  in terms of the  $h_i(x)$  as a rotation in function space? Find an approximate formula for the coefficients in such a transformation. Find a formula for the new expansion coefficients  $\widetilde{A}_i$ .

**Suggestion for Laboratory Excercise.** Use a computer to evaluate

$$1/(2\pi) \int_{-\infty}^{\infty} \exp(-k^2/4n) \exp(ikx) dk = 1/\pi \int_0^{\infty} \cos kx \exp(-k^2/4n) dk ,$$

and compare with  $\sqrt{n/\pi} e^{-nx^2}$ , thereby validating that the FT of a gaussian is again a gaussian. Different values for both  $\delta k$  and  $k_{\max}$  should be tried. Example 1.2 may be useful.

## Further Reading

M.J. Lighthill: *Introduction to Fourier Analysis and Generalized Functions* (Cambridge University Press, 1958).

The mathematical results that we have only stated are found proved in this concise monograph written on an undergraduate level.

R.D. Stuart: *An Introduction to Fourier Analysis*. (Science Paperback, Hals-  
ted, New York 1966).

A more conventional approach to Fourier analysis on an elementary level is found in this concise presentation.

K.F. Riley, M.P. Hobson, S.J. Bance: *Mathematical Methods for Physics and Engineering*. (Cambridge University Press, Cambridge, UK and New York, NY 1998).

Has everything.

F.W. Byron, Jr., R.W. Fuller: *Mathematics of Classical and Quantum Physics*. (Dover Publications, New York, NY 1992).

Hilbert Space is discussed fully here.

H. Margenau, G.M. Murphy: *The Mathematics of Physics and Chemistry*. (Van Nostrand, New York 1943).

A very good alternate exposition of many topics studied here, at about the same level.

A. Erdelyi (ed.): *Tables of Integral Transforms* (McGraw-Hill, New York 1954).

These are, perhaps, the most famous integral transform tables.

## 2. Oscillations of Mechanical and Electrical Systems

**Abstract.** The equation of motion for a mechanical system comprised of a mass at the end of a string and vibrating in a viscous medium is compared to the analogous equation for an electrical system consisting of an R-L-C circuit. The distinction between natural and driven dynamics is carefully drawn. For natural motion the initial value problem is solved for the case of negligible viscous damping, negligible electrical resistance. The resonance phenomenon is brought out fully in connection with the analysis of forced motion, and the characteristics of resonance are related to those of natural motion, attention being drawn to the quality factor  $Q$ . The Green's function method of finding the solution for forced motion is described. A subsequent section is then devoted to an explanation of the operation of oscillators in general, and extended to a description of an actual pendulum clock. In the problems we treat natural motion for nonnegligible damping.

### 2.1 The Systems and Their Equations

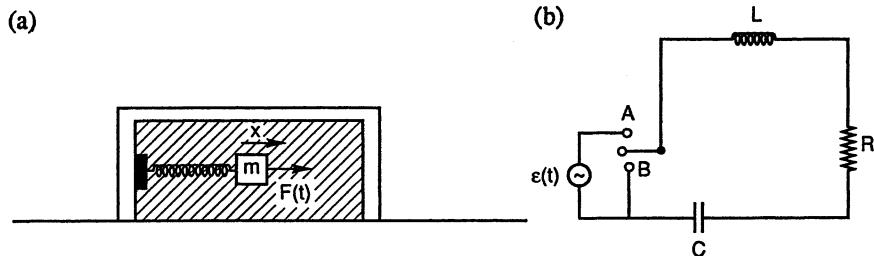
Figure 2.1a shows a mass  $m$ , at the end of a spring of spring constant  $k$ , suspended in a viscous medium which exerts a back force on the mass of  $b\dot{x}/dt$ , with  $\dot{x}$  being the velocity of the mass which has been displaced a distance  $x$  from its equilibrium or resting position. If, further, an *external* time-dependent force  $F(t)$  is applied to the mass, then the total force on the mass will be

$$F(t) - b\frac{dx}{dt} - kx .$$

From Newton's second law,  $F = ma$ , we have

$$F(t) = m\frac{d^2x}{dt^2} + b\frac{dx}{dt} + kx . \quad (2.1)$$

Figure 2.1b shows an electrical system with an inductor  $L$ , a resistor  $R$ , and a capacitor  $C$  all connected in series. If an external time-dependent emf  $\varepsilon(t)$  is applied to the system, then recognizing that a charge going once around the loop does not change its energy, i.e., setting the voltages around the loop equal to zero, we have



**Fig. 2.1.** (a) Linear lumped mechanical system under action of driving force  $F(t)$ . A mass is immersed in a viscous medium. Natural damped oscillations occur when  $F(t) = 0$ . (b) Electrical circuit which is the analogue of (a). With the switch at B, the circuit undergoes natural damped oscillations, assuming the capacitor C was initially charged. With the switch at A the circuit undergoes forced oscillations under the driving influence of the voltage  $\varepsilon(t)$

$$\varepsilon(t) = L \frac{dI}{dt} + RI + \frac{q}{C} \quad (2.2)$$

and since  $I$ , the current, is given by  $dq/dt$ , with  $q$  being the electric charge on the capacitor,

$$\varepsilon(t) = L \frac{d^2q}{dt^2} + R \frac{dq}{dt} + \frac{q}{C}. \quad (2.3)$$

Expressions (2.1) and (2.3) are seen to be the same equation if we identify

$$F(t) \quad \text{with} \quad \varepsilon(t) \quad (2.4a)$$

$$x \quad \text{with} \quad q \quad (2.4b)$$

$$m \quad \text{with} \quad L \quad (2.4c)$$

$$b \quad \text{with} \quad R \quad (2.4d)$$

$$k \quad \text{with} \quad \frac{1}{C}. \quad (2.4e)$$

This means that if we have an appropriate solution to (2.1), then it is also a solution of (2.3), and vice-versa. For this reason we speak of the circuit in Fig. 2.1b as being the electrical analogue of the system shown in Fig. 2.1a, and, similarly, the system in Fig. 2.1a as the mechanical analogue of the circuit of Fig. 2.1b.

The mechanical and electrical quantities listed in (2.4) are, in general, physically quite different. Let us review what we remember about them. All matter has mass, and in classical physics all material objects satisfy Newton's laws. The property of electric inductance,  $L$ , comes about as a result of Ampere's law, according to which electric currents lead to magnetic fields, and Faraday's law, according to which changing magnetic fields induce voltages. The sense of these voltages, according to Lenz's law, is such as to oppose whatever changes in magnetic flux are taking place. The net result is a back

voltage  $LdI/dt$ , where the inductance  $L$  is a measure of the coupling between magnetic flux and circuit current.

Turning to the coefficient  $b$ , perhaps the most familiar cause of viscous drag we have is air resistance (a high speed automobile wastes a great deal of energy in warming up the surrounding atmosphere). The origin of the electrical resistance  $R$  is also frictional. In this case the main impediment to the flow of electric charges, actually electrons, is caused by the irregular thermal vibrations of the atoms comprising the conductor. A back force proportional to the average forward velocity of the electrons results.

The force  $-kx$ , where  $x$  is the displacement from equilibrium, is the familiar Hooke's law for a spring. A spring is not only something found in mattresses or in the laboratory. Every mechanical system acts as a spring as long as its deformation is within its elastic limit. Even the vertical motion of the center of a complex structure such as a suspension bridge stressed by a truck passing over it can be represented by a spring system. The electrical capacitance  $C$  is defined as the ratio of the charge on the capacitor,  $q$ , divided by the voltage required to maintain that charge on the capacitor. We remember that accumulation of electric charge leads to electric fields, and that the effect of these fields is to make it difficult to bring additional charge to the accumulation. A large capacitor is one where much charge leads to small fields, so that even the last charge can be put on the capacitor with a small expenditure of energy. Just as every mechanical system has some resilience, so has every electrical system some capacitance. Not only is there some capacitance between the metal plates of the little air capacitor used to tune radio sets, there is likewise capacitance, for instance, between a high voltage transmission line and the ground beneath it.

Indeed, (2.1) describes every *linear* mechanical system, as long as we are content to regard it as a *lumped* system. Similarly, (2.3) describes every linear lumped electric system. The word "lumped" means that the time dependence of a single parameter contains all the information that we want. For example, the full motion of a suspension bridge includes the relative motion of the different parts of the bridge rather than just the motion at its center. The bridge is in its detailed motion a *distributed* system. In the case of an electric transmission line, if it is long we could never be satisfied with a lumped treatment, because owing to the distributed capacitance between incoming and outgoing conductors, the electrical charge  $q$  would at any instant of time vary as we go down the line. We shall learn in later chapters how the treatment of distributed systems rests on the treatment of lumped systems, to be undertaken here.

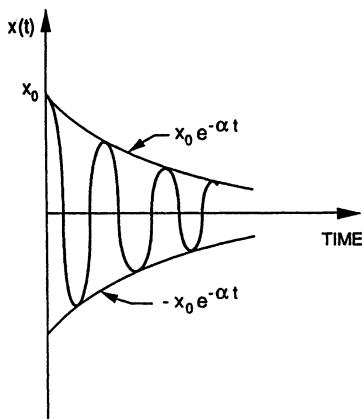
Equation (2.1) is a *linear* differential equation. We note that the displacement and its various derivatives are all raised to the power of unity. If we were to distend the spring beyond its elastic limit the simple linear Hooke's law would break down, and the governing differential equation would no longer be linear. If in the electrical system the current were so large as to cause

the electric resistance to change because of heating of the conductor, i.e., if  $R$  depended on  $dq/dt$ , (2.3) would no longer be linear. Linear differential equations have one extremely valuable property, namely, if we have found a particular solution to the equation for a given driving force, and then go on to find a second solution for a new driving force, and a third solution, then the sum of any two of these or the sum of all of them are likewise solutions of the equation for a driving force which itself is the sum of the component forces. This means that when faced with the problem of solving (2.1) for a rather complicated driving force  $F(t)$  (driving forces can be complicated in practice), we can resort to the Fourier representation discussed in Chap. 1. We can write  $F(t)$  as the sum of components, each component having only one simple harmonic motion in time, solve the problem for each of these component driving forces, and then add together all the resulting solutions. See the Problems for illustration.

The next characterization of (2.1, 3) which we need to understand is striking, and of paramount importance. It is the distinction whether we do, or do not, have a nonvanishing left side; that is, whether we do or do not apply a driving force  $F(t)$  to the mechanical system, or a driving emf,  $\varepsilon(t)$ , to the electrical system. If we do not, the differential equations are said to be homogeneous. If  $F(t)$  is zero, then the mass, after being released with some initial displacement or velocity, or both, will move at the will of the system. We call this free motion of the system its *natural* motion. Similarly we can remove the driving voltage  $\varepsilon(t)$  of our electrical system by switching to position B, Fig. 2.1b, thereby obtaining its natural response. This problem is different from the situation when a driving force  $F(t)$ , or a driving voltage  $\varepsilon(t)$  exist. For example, someone could be hitting the mass in Fig. 2.1a with a hammer on a regular basis, and clearly then  $x(t)$  would be different from the natural motion. The solution of (2.1) with the given  $F(t)$  describes the *forced* motion of the mass. Similarly, if we switch back to A in Fig. 2.1b then the time dependence of the charges and currents will after a short time be exclusively influenced by the electrical voltage  $\varepsilon(t)$  of the generator, a forced response of the electric circuit. The linear nature of the governing differential equations gives the general solution as the sum of the forced solution and the natural solution, the latter determined by the initial conditions. We shall see that in practice the natural solution becomes damped out and is referred to as the transient. Accordingly, we shall divide our subsequent work into two sections. In the first we shall analyze the natural motion of the systems, in the second their forced responses. Moreover, as we shall see, there are important relationships between the two responses.

## 2.2 Natural Motion of Systems

We know from experience what happens when the mass in Fig. 2.1a is displaced from its equilibrium position at  $x = 0$ , moved to some point  $x = x_0$ ,



**Fig. 2.2.** On the basis of observation we expect the mass of Fig. 2.1 to oscillate with ever smaller amplitude when no external force is applied other than air resistance. For a  $Q$  of 2000, we would have to show about 600 oscillations in our figure, which we have extended over a time  $1/\alpha$

and released, let us say, with no starting velocity. The mass accelerates as it is pulled back to the equilibrium position by the spring. When it reaches  $x = 0$  the force exerted on it by the spring vanishes. However, the mass, which at that point has considerable kinetic energy, a maximum as it turns out, and would require an infinite force to stop it quite suddenly, continues its motion until it reaches close to  $x = -x_0$ . At this point its kinetic energy has been reabsorbed by the spring, the mass comes to a stop, and begins its motion in the reverse direction. We say "close" to  $x = -x_0$  because we know that a small amount of energy has been lost to the viscous damping. The mass, accordingly, undergoes a motion as in Fig. 2.2. We refer to this motion as *damped oscillations*.

Analytically, damped oscillations have the form

$$x(t) = (A' \sin \omega t + B' \cos \omega t)e^{-\alpha t}, \quad (2.5)$$

i.e., the usual form for oscillations of angular frequency  $\omega$  here multiplied by the factor  $\exp(-\alpha t)$ , with  $\alpha$  a positive coefficient. The factor  $\exp(-\alpha t)$  gets smaller in a smooth way with time, and will produce a damping of the oscillations. The coefficients  $A', B'$  are amplitudes; their size has to do with the beginning displacement and velocity of the mass, questions we put off for the moment.

To check (2.5) and see what are the quantities  $\omega$  and  $\alpha$ , we substitute  $x(t)$  into the equation of motion (2.1), with the driving force set equal to zero:

$$m \frac{d^2x}{dt^2} + b \frac{dx}{dt} + kx = 0. \quad (2.6)$$

To make things less tedious let us write

$$x(t) = A e^{i\omega t} e^{-\alpha t} = A e^{i\Omega t} \quad (2.7a)$$

$$\Omega = w + i\alpha. \quad (2.7b)$$

Here  $A$  is a complex number, and we have written the exponent in terms of a “complex frequency”  $\Omega$ . We remember from Chap. 1 that by a judicious choice of phase for the complex number  $A$ , the real part of  $A \exp(i\omega t)$  can be made into any desired combination of sines and cosines (Problem 1.9). At the same time  $x(t)$ , if defined as in (2.7a), gains an imaginary part. The linear equation of motion with real coefficients will determine independently both the real and imaginary parts of  $x(t)$ . But  $x(t)$  describes the motion of a mass, how can it be a complex number? The answer is that although we work with  $x(t)$  complex, the real part of  $x$  is the physical solution of whatever differential equation we are working with. The imaginary part is excess baggage, which we discard in the end. While this sounds a bit complicated, it actually is very practical, as we shall now see, mostly because exponential functions are so easy to differentiate.

We substitute (2.7a) into (2.6), and perform the indicated differentiations:

$$m(i\Omega)^2 A e^{i\Omega t} + b(i\Omega) A e^{i\Omega t} + k A e^{i\Omega t} = 0 .$$

We see a common factor,  $A \exp(i\Omega t)$ , meaning that a solution of the form (2.7a) will, indeed, satisfy (2.6). Cancelling this factor we find  $\Omega$  from

$$-m\Omega^2 + i\Omega b + k = 0$$

which upon substituting from (2.7b) becomes

$$-m(\omega^2 - \alpha^2 + 2\omega i\alpha) + ib(\omega + i\alpha) + k = 0 .$$

Separating real and imaginary parts we get

$$\begin{aligned} -m\omega^2 + m\alpha^2 - \alpha b + k &= 0 \\ -2\omega\alpha m + b\omega &= 0 . \end{aligned}$$

We find

$$\alpha = \frac{b}{2m} \tag{2.8a}$$

from the second equation, and from the first,

$$\begin{aligned} -m\omega^2 + \frac{b^2}{4m} - \frac{b^2}{2m} + k &= 0, \\ \omega &= \sqrt{\frac{k}{m} - \frac{b^2}{4m^2}} . \end{aligned} \tag{2.8b}$$

If we set the viscous damping coefficient  $b$  in (2.8b) equal to zero, we obtain, as we should, the angular frequency for a simple undamped mechanical oscillator of mass  $m$  and spring constant  $k$ ,

$$\omega_0 = \sqrt{k/m} . \tag{2.9}$$

In many applications the frictional term  $b/2m$  in (2.8b) leads to only a small correction to the angular frequency  $\omega_0$ . The angular frequency  $\omega_0$  is called

the *natural frequency* of the system, even though when the system undergoes its natural (unforced), motion the actual frequency is the slightly shifted quantity  $\omega$ . In conclusion, we write (2.8a,b) as

$$\omega_f = (\omega_0^2 - \alpha^2)^{1/2} \quad (2.10)$$

$$\alpha = \frac{b}{2m}. \quad (2.8a)$$

The subscript f emphasizes that we have free oscillations.

Let us go back and find the undetermined quantities  $A'$ ,  $B'$  in the original equation (2.5). We shall do this for the particular problem already posed, namely, at time  $t = 0$  the mass is displaced to the point  $x = x_0$  and released with zero initial velocity, and we shall assume for the present that the frictional term is small. At time  $t = 0$  we have from (2.5)

$$x|_{t=0} = B'. \quad (2.11a)$$

Let us also differentiate (2.5) to find an expression for the velocity of the mass,  $dx/dt$ :

$$\frac{dx}{dt} = \omega_f(A' \cos \omega_f t - B' \sin \omega_f t)e^{-\alpha t} + \text{small term}. \quad (2.12)$$

The small term arises from differentiating the factor  $\exp(-\alpha t)$ , and is small, by virtue of its proportionality to  $\alpha$ , as compared to  $\omega$  in the term which is not small. Leaving out the small term, we find that at  $t = 0$  this equation reduces to

$$\left. \frac{dx}{dt} \right|_{t=0} = \omega_f A'. \quad (2.11b)$$

Thus, for our case of small damping:

$$A' = 0 \quad (2.13a)$$

$$B' = x_0 \quad (2.13b)$$

$$x(t) = x_0 e^{-\alpha t} \cos \omega_f t. \quad (2.13c)$$

Evidently, we can fit the amplitudes  $A'$ ,  $B'$  to whatever initial combination of displacements and velocities the mass acquired at time  $t = 0$ . In Problem 2.5 we solve the initial value problem exactly with a finite  $\alpha$ . However, the case of negligible damping shall appear frequently in later chapters. We have already drawn attention to Fig. 2.2 showing  $x(t)$ . The two curves  $\pm x_0 \exp(-\alpha t)$  act as envelopes. Every time the cosine takes on the value +1, i.e., at time intervals of  $2\pi/\omega_f$ ,  $x(t)$  touches the top curve; similarly, at the half periods, when  $\cos \omega_f t$  is -1,  $x(t)$  touches the lower curve. This is true even if  $\cos \omega t$  contains a phase shift, as it does for finite  $\alpha$ . Note that the points of contact are not actually the maxima or minima of the cycles (Problem 2.4).

Having solved the problem of the natural motion of the mechanical system, we have the solution of the analogous problem for the electrical system

likewise in hand. All we need to do is to use the identifications listed in the relationships in (2.4). We see that in the electrical problem at  $t = 0$  the capacitor has some given initial charge, say  $q_0$ , while the initial current in the circuit vanishes,  $dq/dt = 0$ . By analogy with (2.13c) the subsequent time development of the charge on the capacitor is given by

$$q(t) = q_0 e^{-\alpha t} \cos \omega_f t . \quad (2.14)$$

Here, in analogy with (2.8–10),

$$\omega_f = (\omega_0^2 - \alpha^2)^{1/2} . \quad (2.15a)$$

The natural frequency of the system, again  $\omega_0$ , is given by

$$\omega_0 = \sqrt{1/LC} , \quad \text{and} \quad (2.15b)$$

$$\alpha = R/2L . \quad (2.15c)$$

It also follows that Fig. 2.2 is a representation of  $q(t)$  as well as of  $x(t)$ . Finally, we note that whereas in the mechanical system the energy shifts back and forth between the spring (when the displacement is at a positive or negative peak) and the kinetic energy (when the velocity is at a peak), in the electrical system the energy is shifting from the capacitor when the charge is at a peak to the inductor when the current is at its peak.

While the mechanical and electrical systems are qualitatively similar in so many aspects, a striking quantitative difference exists in the number of cycles the system will undergo before the motion becomes damped out. The damping,  $\exp(-\alpha t)$ , suggests that  $1/\alpha$  is a measure of the time for which the system can oscillate, and is commonly referred to as the time constant. Since the period for a cycle is very nearly given by  $2\pi/\omega_0$ , the ratio  $(1/\alpha)/(2\pi/\omega_0)$  is some measure of the number of cycles allowed. For an electrical system we have

$$\text{ratio} = \frac{2L/R}{2\pi/\omega_0} = \frac{\omega_0 L}{\pi R} \quad (2.16a)$$

and for a mechanical system

$$\text{ratio} = \frac{2m/b}{2\pi/\omega_0} = \frac{\omega_0 m}{\pi b} . \quad (2.16b)$$

It is conventional to define a quality factor  $Q$  given by

$$Q = \omega_0 L/R \quad \text{electrical} ; \quad (2.17a)$$

$$Q = \omega_0 m/b \quad \text{mechanical} . \quad (2.17b)$$

$Q$  tends to be higher for electrical than for mechanical systems, indicating the greater number of oscillations sustained electrically. While for a mechanical system a  $Q$  of 50 is not unreasonable, electrical systems working at radio frequencies may have  $Q$ s of the order of several thousand.

## 2.3 Forced Motion

As mentioned, the driving force  $F(t)$  appearing on the left of (2.1) is generally not a simple function. Our scheme here will be to find the motion  $x(t)$  corresponding to a single arbitrary Fourier component of that force, i.e., a simple harmonic time dependence. If one wishes, one can then add the solutions for all the different Fourier components present in the driving force to get the overall response. We thus write

$$F(t) = F_0 \cos \omega t . \quad (2.18a)$$

In the case of the analogous electrical problem the solution with a harmonic drive is often in itself of practical use; many electric generators are designed to produce pure harmonic voltages of adjustable or fixed frequency  $\omega$ :

$$\varepsilon(t) = \varepsilon_0 \cos \omega t . \quad (2.18b)$$

Combining (2.18a) with (2.1), we see that the problem we wish to solve is contained in the differential equation

$$m \frac{d^2x}{dt^2} + b \frac{dx}{dt} + kx = F_0 \cos \omega t . \quad (2.19)$$

Let us attempt to solve this equation, again as in the previous section, by letting  $x(t)$  become a complex quantity, and discarding its imaginary part after finding the solution. Since imaginary parts are to be discarded, it is convenient to begin by rewriting (2.19) in the form

$$m \frac{d^2x}{dt^2} + b \frac{dx}{dt} + kx = F_0 e^{i\omega t} . \quad (2.20)$$

The real part of this equation is just (2.19). Similarly, we write

$$x(t) = A e^{i\omega t} \quad (2.21)$$

where  $A$  is an unknown constant, a complex quantity to allow for phase lag which we wish to find. We notice that  $x(t)$  is assumed to have precisely the same angular frequency as the driving force,  $\omega$ . It is an experimental fact, which also follows from the orthogonality of Fourier waves, that if a system is steadily driven at one frequency, it will respond at the same frequency.

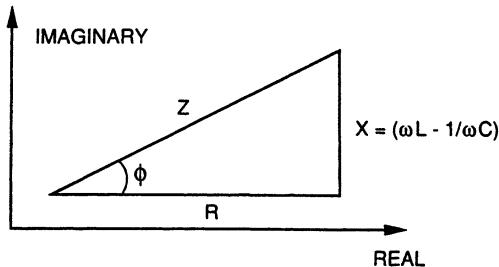
Let us substitute (2.21) into (2.20) and do the indicated differentiation. We find

$$(i\omega)^2 m A e^{i\omega t} + i\omega b A e^{i\omega t} + k A e^{i\omega t} = F_0 e^{i\omega t} .$$

Again we have a common factor  $e^{i\omega t}$ , which we cancel, and solve for  $A$ :

$$\begin{aligned} A &= F_0 / (-m\omega^2 + i\omega b + k) \\ &= F_0 (-m\omega^2 + k - i\omega b) / [(-m\omega^2 + k)^2 + \omega^2 b^2] . \end{aligned} \quad (2.22)$$

In the last equation we have written  $A$  as the sum of a real and an imaginary component.



**Fig. 2.3.** The positive phase angle  $\phi$

We can simplify this result by introducing the mechanical *reactance*  $X$ , where

$$\omega X = m(\omega^2 - k/m) = m(\omega^2 - \omega_0^2). \quad (2.23a)$$

We find from (2.22):

$$A = -\frac{F_0}{\omega} \left( \frac{X + ib}{X^2 + b^2} \right) = \frac{Z^* F_0}{i\omega |Z|^2}$$

where  $Z$ , the complex mechanical *impedance*, is defined by

$$Z = b + iX. \quad (2.23b)$$

The triangle in the complex plane in Fig. 2.3 identifies a positive *phase angle*  $\phi$ . We can therefore write

$$Z = |Z|e^{i\phi}$$

and

$$A = \frac{-iF_0 e^{-i\phi}}{\omega |Z|}.$$

It follows that

$$\begin{aligned} x(t) &= \operatorname{Re} \{ A e^{i\omega t} \} \\ &= \operatorname{Re} \left\{ \frac{F_0 \exp i(-\phi + \omega t - \pi/2)}{\omega |Z|} \right\} \end{aligned}$$

since  $-i$  is also  $\exp(-i\pi/2)$ . Or we have

$$x(t) = F_0 \sin(\omega t - \phi) / \omega |Z| \quad (2.24a)$$

where

$$\sin \phi = X / |Z| \quad (2.24b)$$

and

$$|Z| = [b^2 + X^2]^{1/2}. \quad (2.24c)$$

Before discussing this result, let us go on to the electrical analogue problem, and write down the corresponding result when the electrical system is driven by the simple harmonic voltage  $\varepsilon(t)$  given in (2.18b). Rather than giving the results in terms of the charge on the capacitor  $q$ , it is customary to give them in terms of the electric current  $I = dq/dt$ , that is, the quantity which is the analogue of the velocity  $dx/dt$  in the mechanical problem. We therefore differentiate (2.24a) with respect to the time, and also do the replacements of (2.4). We thus find for the electric current

$$I(t) = \varepsilon_0 \cos(\omega t - \phi)/|Z| \quad (2.25a)$$

with

$$Z = R + iX, \quad (2.25b)$$

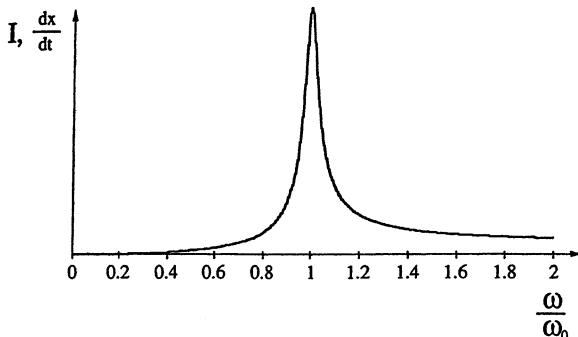
$$X = \omega L - \frac{1}{\omega C} = \frac{L}{\omega}(\omega^2 - \omega_0^2), \quad (2.25c)$$

and  $\phi$  is again  $\sin^{-1}(X/|Z|)$ .

What chiefly interests us about this result is the way the current changes when we fix the amplitude of the driving voltage  $\varepsilon_0$  and vary the driving angular frequency,  $\omega$ . This is commonly referred to as the *response* of the system. To see this response most clearly we plot in Fig. 2.4 a graph of current amplitude versus driving frequency  $\omega$ , for fixed input voltage  $\varepsilon_0$ . Since  $R$  is generally a rather small quantity the graph will go to a sharp maximum at

$$\omega = \omega_0. \quad (2.26)$$

Here the frequency dependent term in the denominator of (2.25a), i.e. in  $Z$ , which is never negative, vanishes. Further, we see from (2.24c) that in the immediate neighborhood of  $\omega_0$  the response curve will decrease in a smooth and almost symmetric manner.



**Fig. 2.4.** Resonance response, the same function for both electrical and mechanical systems. Electric current, or particle velocity, plotted as a function of the angular frequency of the applied drive, with the amplitude of the drive maintained constant. Resonance (optimum response), takes place at the natural frequency of the systems,  $\omega_0$ . The plot is for  $Q = 20$

The condition (2.26) is called *resonance*, a word which implies what actually happens at this frequency, maximum response for a fixed input. Indeed, as  $R$  goes to zero the amplitude tends to infinity. From (2.24) at resonance the phase shift  $\phi = 0$  for all  $R$ ;  $\epsilon_0(t)$  and  $I(t)$  are in phase; the circuit acts as a pure resistor, and positive power flows into it at all times. For fixed driving voltage  $\epsilon_0$  the absorbed power is a maximum (Problem 2.8). Off resonance  $\phi \neq 0$ , and when  $R = 0$  power going into the circuit during part of a cycle is given back during the remainder of the cycle. However, at resonance if  $R$  vanishes the amplitude builds since power is then continually absorbed by the oscillator without dissipation.

Mechanical systems have responses that show resonance just as electrical systems do. Anyone who has ever watched a heavy vehicle being “rocked” from a stranded position in a slippery depression, or who has seen young ladies and gentlemen riding a ski chair-lift and finding it fit to rock with just the right timing until all the chairs undergo vertical motions of perilous amplitude has witnessed precisely mechanical resonance. Many of our daily experiences of resonance are in connection with distributed systems, of which the details are left to ensuing chapters.

The electric system is one of the supporting pillars of electronic technology. Communication receivers have such a circuit as their first stage, which is adjusted to resonate at (is tuned to) the frequency of the desired incoming signal. Second, the circuit is a basic component of a frequently employed type of oscillator (L–C oscillator). The operation of most communication systems is dependent on stable oscillators in both transmitters and receivers.

## 2.4 Forced Motion and the Green’s Function

There is another way of solving (2.1) for the forced motion aside from the Fourier representation method described in the previous section – the Green’s function method, after the English physicist George Green (1793–1841). The latter method with its clear physical motivation is, when the relevant Green’s function is available, more direct, and the results are often easier to understand. Moreover, as we shall see in Chap. 6, the method can also be carried over into quantum mechanics, where the Green’s function plays a significant role.

The basic idea is as follows: We wish to solve the forced motion equation:

$$m \frac{d^2x}{dt^2} + b \frac{dx}{dt} + kx = F(t) . \quad (2.1)$$

Instead, we first solve the equation:

$$m \frac{d^2G(t - t_0)}{dt^2} + b \frac{dG(t - t_0)}{dt} + kG(t - t_0) = \delta(t - t_0) . \quad (2.27)$$

That is, the driving force  $F(t)$  is replaced by a single impact at time  $t_0$  with an impulse of unity. We, then, represent  $F(t)$  by a sum of such impacts. The

response to the impact at  $t_0$  is  $G(t - t_0)$ . We can expect on physical grounds that the total response  $x(t)$  will be given by the sum of all the responses. that is,  $x(t)$  will be given by:

$$x(t) = \int_{-\infty}^t F(t_0)G(t - t_0)dt_0 . \quad (2.28)$$

Here  $G(t - t_0)$  is the Green's function, with  $t_0$  the time of impact. It has been assumed that for negative time, i.e.  $t < t_0$ ,  $G(t - t_0)$  vanishes, and that  $F(t_0)$  is defined over all time  $t_0; \infty > t_0 > -\infty$ .

We can give an analytic derivation of (2.28). We are interested in all possible times of impact  $t_0$ . Hence, we are interested in  $\int_{-\infty}^{\infty} F(t_0)\delta(t - t_0)dt_0$  on the right or driving side of (2.27), and in any arbitrary time  $t$  on the left of the same equation. If we multiply both sides of (2.27) by  $F(t_0)$ , and integrate over  $t_0$  we find, interchanging differentiation and integration:

$$\begin{aligned} m \frac{d^2}{dt^2} \int_{-\infty}^{\infty} F(t_0)G(t - t_0)dt_0 + b \frac{d}{dt} \int_{-\infty}^{\infty} F(t_0)G(t - t_0)dt_0 \\ + k \int_{-\infty}^{\infty} F(t_0)G(t - t_0)dt_0 = \int_{-\infty}^{\infty} F(t_0)\delta(t - t_0)dt_0 \end{aligned} \quad (2.29a)$$

$$= F(t) . \quad (2.29b)$$

Again, recalling

$$G(t - t_0) = 0 \quad \text{for } t \leq t_0 \quad (2.30a)$$

we see that  $x(t)$  as given in (2.28) is, indeed, the desired solution. Evidently, future values of  $F(t_0)$ , ( $t_0 > t$ ), do not affect  $x(t)$ .

The point about this method is that, in general, the Green's functions are rather easy to find. We must bear in mind that while these functions describe the response of the system to some sudden blow, they are, after all, describing unperturbed motion, i.e., natural motion. Accordingly, except at the starting time  $t = t_0$ ,  $G(t - t_0)$  satisfies the homogeneous equation of motion (2.6), which, as in 2.2 restricts the solution to damped sines or cosines. Now,  $G(t - t_0)$  is considered zero prior to the impact, the oscillator is standing still, and it can not with a finite force acquire a finite displacement in vanishing time. We must start with vanishing  $G(0)$  as in (2.30a). This rules out any cosine contribution. Thus, we find that:

$$G(t - t_0) = 0 \quad \text{for } t \leq t_0 \quad (2.30a)$$

$$= e^{-\alpha(t-t_0)} \frac{\sin \omega_f(t - t_0)}{\omega_f m} \quad t \leq t_0 \quad (2.30b)$$

$$\omega_f^2 = \omega_0^2 - \alpha^2 . \quad (2.10)$$

The amplitude in (2.30b) comes from the second initial condition, which requires that the initial momentum of the oscillator equal the impulse imparted at  $t = t_0$ :

$$\int_{-\infty}^{\infty} F dt = \int_{-\infty}^{\infty} \delta(t - t_0) dt = 1 . \quad (2.31a)$$

Checking, we have from (2.30b):

$$\begin{aligned} p(t) \Big|_{t=t_0} &= m \frac{\partial G(t - t_0)}{\partial t} \Big|_{t=t_0} \\ &= e^{-\alpha(t-t_0)} \left[ \cos \omega_f(t - t_0) - \frac{\alpha}{\omega_f} \sin \omega_f(t - t_0) \right] \Big|_{t=t_0} . \end{aligned} \quad (2.31b)$$

The final result for  $x(t)$  is (2.28) with  $G(t - t_0)$  given by (2.30a,b). This result must be consistent with our previous derivation using a harmonic driving force  $F_0 \cos(\omega t)$  as given in (2.18a). The interested reader is referred to Problem 2.9. There it is shown that the Fourier transform  $G(\omega)$  of  $G(t)$  is given by:

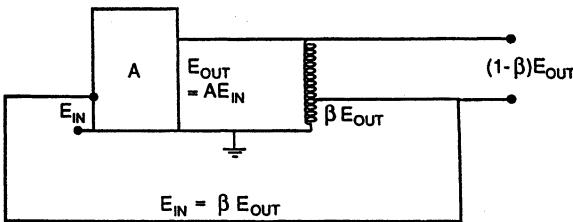
$$G(\omega) = \frac{-1}{\sqrt{2\pi} m(\omega^2 - 2i\alpha\omega - \omega_0^2)} . \quad (2.32)$$

Evidently,  $G(\omega)$  shows the same resonance response as obtained previously.

We shall have occasion to apply the Green's function method to forced motion problems in classical physics in each of the subsequent three chapters. It turns out that the Green's function is likewise important in quantum mechanics. Indeed, as we shall see in Chap. 6, it is one of the bridges taking us from classical mechanics to ordinary wave mechanics to quantum field theory.

## 2.5 Oscillators

Let us look briefly at how an L-C oscillator works. Its central component is an amplifier, a device that enlarges incoming signals by a factor  $A$ , taking the necessary energy from elsewhere. A portion  $\beta$  of the output of this amplifier is fed back to become its input, Fig. 2.5. The portion  $(1 - \beta)$  of the output of the amplifier is available for useful outside application. To begin with, the reader must imagine a variety of meaningless electrical signals going around the loop (noise); their origin could be thermal or, perhaps, there are stray signals picked up from other electrical devices. We can always take the Fourier time transform of all this disturbance. All possible frequencies will be represented. The quantities  $A$  and  $\beta$  are frequency-dependent, and complex to allow for phase shifts. Now, because of the vastly enhanced response of the L-C circuit incorporated (into the load) in the amplifier near  $\omega_0$ , the amplification factor



**Fig. 2.5.** Operation of an oscillator. We see that  $A\beta = 1$  for stable oscillation

$A$  will be large and particularly sensitive to frequency in the neighborhood of  $\omega_0$ . Thus, it will be that for some particular  $\omega$  somewhere near resonance, the conditions for growth,  $\beta A$  real and

$$\beta(\omega)A(\omega) > 1 \quad (2.33)$$

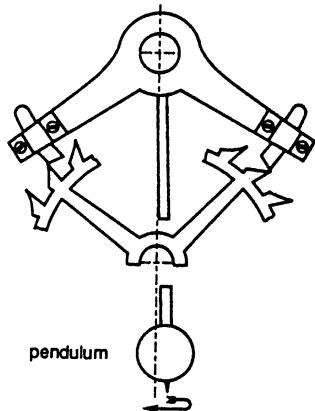
will be satisfied. The condition (2.33) tends to produce growth because each time the signal goes around its loop it increases. What actually happens is that as soon as the system is turned on, nonlinear effects limit  $A$  to where

$$\beta(\omega)A(\omega) = 1. \quad (2.34)$$

The point to emphasize is that (2.33, 2.34) involve phase as well as magnitude. A complex  $\beta A$  implies destructive interference rather than growth. The phase shifts carried by  $A$  and  $\beta$  must be equal and opposite. We can see from (2.24b) that in the vicinity of resonance not only does the magnitude of the response have its most striking change, but the phase does as well. This in turn means that (2.34) will only hold for a rather well-defined frequency, or that the output of the oscillator will be steady in frequency, i.e. stable, the basic requirement for communication systems.

In Example 2.1 we will show an additional property of the resonating systems, namely that the higher the value of the quality factor  $Q$  (2.17a,b), the narrower, more discriminating the resonance curve gets, and the faster will the phase change be as we go through resonance. This fact allows us to link frequency stability with  $Q$ , and accounts for engineers' constantly seeking ways of obtaining higher  $Q$ s. Lastly, we illustrate how the described scheme of oscillator operation applies to a mechanical clock driven by a pendulum which, unlike the electrical case, we can do in detail.

Figure 2.6 shows the *escapement*. Attached to the pendulum is a release lever with two pawls. In view of its shape the lever is referred to as the anchor piece. It controls the motion of the ratchet below it, which itself is under tension from the main spring of the clock. If unimpeded by the anchor the ratchet would spin clockwise. As shown, the pendulum bob is moving to the right. On its swing back, just to the left of the position shown, the left anchor pawl will release the ratchet, which will then advance about a tooth width, being finally held back by the right anchor pawl from advancing further. "Tick" goes the clock. The reverse happens as the bob approaches center

**Fig. 2.6.** Escapement of a pendulum clock

from the left, when the right tooth releases the ratchet – “tock”. The ratchet has advanced one tooth and slot, the pendulum has gone through one cycle.

Essential is that energy be supplied to the pendulum without greatly affecting its period. Indeed, the slopes of the pawl tips are such as to assure the flow of energy from ratchet to anchor, i.e. to pendulum. To understand the representation of the clock as an amplifier with feedback, one can imagine replacing the ratchet by a periodic drive of the pendulum and anchor. We know from our analysis of driven motion that for a fixed driving force we will secure a large amplitude only near the natural frequency. That is, only here would the motion of the anchor be large enough to control a ratchet. In reality the ratchet is feeding energy back to the (air) damped pendulum near resonance and at the proper phase. Concurrently, it is providing output, driving the gearing which connects it to the clock hands. Thus, if the pendulum is initially driven, or simply released, it will have to settle down to oscillate at an  $\omega$  very near its natural frequency, where  $\beta(\omega)A(\omega)$  can equal unity. (The amplifier consists of the pendulum, the escapement, and the power source, its output being in the motion of the ratchet. The signal fed back from ratchet to anchor gets amplified.)

## 2.6 Summary

We have found that both the mechanical system and its electric analogue, if excited and then left to themselves, will undergo oscillations that become damped at a rate dependent on the magnitude of the frictional agent involved, be that viscous drag or electric resistance. The angular frequencies of these oscillations are only slightly shifted from the values they would assume in the absence of any damping, the so-called natural frequencies of the systems. When these same systems are driven by inputs having a simple harmonic time dependence, they exhibit a sharply increased response when the driving

frequencies become equal to the natural frequencies that characterize the free motion. By “response” we mean the change in amplitude of the velocity of the mass or of the electric current with driving frequency, and the phenomenon is called resonance.

We also note that systems with high quality factors  $Q$ , on the one hand undergo many oscillations before they are damped when the systems undergo their natural motion, and on the other are those systems with particularly sharp resonance responses. The connection between these facts and the use of the systems in oscillators and clocks is not entirely direct, originating with the rapid changes in phase near resonance with frequency. It is the distinctiveness of the resonance phenomena that leads to steady oscillation. This is ultimately what one expects on an intuitive basis.

**Example 2.1.** Show that the width of resonance curves can be directly related to quality factors  $Q$ .

*Solution.* By setting the two terms in the denominator of (2.25a) equal to each other we will find a condition on the driving frequency which reduces the current to  $1/\sqrt{2}$  of its value at resonance, a so-called half-power point. Taking the square root of both terms, we find

$$L(\omega_0 - \omega) \frac{(\omega_0 + \omega)}{\omega} = R .$$

Since we are still in the immediate vicinity of resonance we replace  $\omega$  by  $\omega_0$  in the third factor on the left (but not the second!) and write

$$L(\omega_0 - \omega)/2 \cong R .$$

For the width of the curve at the half-power point,  $\Delta\omega$ ,

$$\Delta\omega = 2(\omega_0 - \omega)$$

so that  $\Delta\omega/\omega_0 = R/\omega_0 L = 1/Q$ .

This is the desired result, the relative width at half-power is inversely proportional to the quality factor  $Q$ .

**Example 2.2.** An electric circuit consisting of an  $R$  and  $L$  and a  $C$  in series is charge-free and inert when suddenly, at  $t = 0$  it is connected across a battery of constant voltage  $V_0$ . Find the subsequent time dependence of the charge on the capacitor,  $q(t)$ . Consider  $R$  small.

*Solution.* The problem can be understood readily via its mechanical analogue. We know that if we suddenly apply a constant force  $F_0$  to the mechanical system, the mass will not move at all in that instant, but the actual equilibrium point of the spring will be displaced forward a distance  $x_0 = F_0/k$ . We have the same situation as in the example, namely a mass is started with zero initial velocity, but with initial displacement, here  $-F_0/k$ , relative to the equilibrium position. Translating back to the electrical system, the answer is

$$q = -CV_0 \cos \omega_f t e^{-\alpha t} + CV_0, \quad \alpha = R/2L .$$

## Problems

**2.1** Combine the laws of Ampere and Faraday to show that the inductance of a long solenoid is given by

$$L = \mu_0 n^2 l A ,$$

where  $n$  is the number of turns per unit length,  $l$  is the total length of the solenoid, and  $A$  its cross-sectional area.

**2.2** Show that if  $x_1(t)$  is the solution of (2.1) when the driving force is  $F_1(t)$ , and  $x_2(t)$  is the solution when  $F_2(t)$  is the driving force,  $(x_1 + x_2)$  is the solution when the driving force is given by  $(F_1 + F_2)$ . What statement can you make for so-called homogeneous differential equations, where the driving forces are zero? What determines  $x(t)$  in this case?

**2.3**  $\alpha = 0$ . When all the charge has been conducted away from the capacitor, why does the electric current in a circuit undergoing its natural motion not come to a stop? You may use the mechanical analogue for help in answering. Assuming that the resistance in the circuit is zero, what is  $dI/dt$  at the instant when  $q = 0$ ?

**2.4**  $\alpha$  small. The previous problem was rather easy. Now discuss the condition

$$dI/dt = 0$$

when the resistance in the circuit, while small, is nonvanishing.

a) Let the time dependence of the current  $I$  be given by:

$$I(t) = I_0(\cos \omega_f t)e^{-\alpha t}, \quad \alpha = R/2L .$$

Draw a graph of  $I(t)$  vs.  $t$ . Do the minima of the curve come before or after the minima of  $\cos \omega_f t$ ? The maxima? Give a simple explanation for your answers.

b) Find analytically the location of the extrema of  $I(t)$ , and check your answers in Part (a).

c) Use the differential equation of motion for  $Q(t)$  to find the charge  $Q$  at the extrema of  $I(t)$ , i.e. when  $dI/dt = 0$ . What is the physical reason that  $Q$  does not vanish as it did for  $\alpha = 0$ ; what will be the sign of  $Q$  compared to that of  $I$ ? Finally, show that at the extrema of  $I(t)$ , to first order in  $\alpha$ ,  $Q(t)$  is given by:

$$Q(t) \cong -(2\alpha/\omega_0^2)I_0e^{-\alpha t} .$$

d) Discuss the mechanical analogue solutions to the questions posed in Parts (a) and (c).

**2.5** a)  $\alpha$  not small. In this case we must use (2.12) rather than (2.11b). Show that the initial conditions

$$q|_{t=0} = Q_0 \quad I|_{t=0} = 0$$

are exactly satisfied if

$$I = -Q_0 \frac{\omega_0^2}{\omega_f} e^{-\alpha t} \sin \omega_f t .$$

*Hint:* Show first that

$$q = Q_0 (\cos \omega_f t + \frac{\alpha}{\omega_f} \sin \omega_f t) e^{-\alpha t} .$$

b) A special case is where  $\alpha$  is so large that

$$\omega_f^2 = \omega_0^2 - \alpha^2 = 0 .$$

This is called critical damping. Derive a simple expression for  $I$  as a function of time valid for critical damping. Draw a rough graph of  $I(t)$  for this case.

*Hint:*  $\lim_{\omega \rightarrow 0} (\sin \omega t)/\omega = t$ .

**2.6**  $\alpha$  large. a) If  $\alpha \gg \omega_0$  we can define two positive decay constants  $\alpha_1, \alpha_2$

$$\alpha_1 = \alpha + (\alpha^2 - \omega_0^2)^{1/2}$$

$$\alpha_2 = \alpha - (\alpha^2 - \omega_0^2)^{1/2} .$$

Show that for the initial conditions in Problem 2.5,  $q(t)$  will be given by

$$q(t) = Q_0 \frac{\alpha_1 e^{-\alpha_2 t} - \alpha_2 e^{-\alpha_1 t}}{\alpha_1 - \alpha_2}$$

for a free oscillator, and that after a long time

$$q(t) \approx Q_0 e^{-t/\tau}$$

where  $\tau$ , the relaxation time, is given by

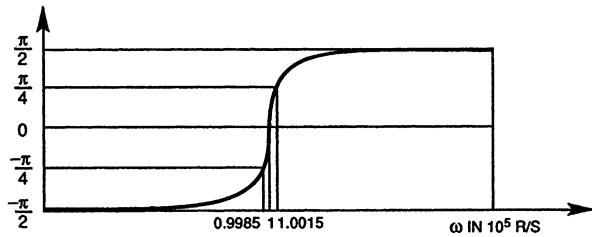
$$\tau = 2\alpha/\omega_0^2 .$$

b) Show that if the oscillator is *driven* at a low frequency  $\omega \ll \omega_0$ ,

$$q(t) = \frac{\varepsilon_0 C}{(1 + \omega^2 \tau^2)^{1/2}} \sin(\omega t - \phi) ,$$

where  $\tau$  is the relaxation time of part (a). What value does  $\phi$  approach as  $\omega \rightarrow 0$ ? What value does the amplitude of the mechanical oscillator approach? Sketch the amplitude of  $x(t)$  as a function of  $\omega$ .

**2.7** a) Verify the statement in the text: "In the vicinity of resonance ... not only does the magnitude of the response have its most striking change, but the phase does as well."



**Fig. 2.7.** The phase  $\phi$  as a function of  $\omega$  near resonance, Problem 2.7b

b) When testing for the resonance response of an L–C–R circuit one found that the phase lag of the current  $\phi$  followed the graph shown in Fig. 2.7. Calculate the number of free oscillations during the time an initial charge  $Q_0$  decays to  $Q_0/e$ .

**2.8** a) Show that the average power absorbed by a mechanical oscillator is given by

$$\bar{P} = \frac{F_0^2}{2|Z|} \cos \phi .$$

b) Verify that the points on the resonance curve selected in Example 2.1 indeed correspond to half-power.

**2.9** a) The solution of the forced oscillator equation

$$m \frac{d^2x}{dt^2} + b \frac{dx}{dt} + kx = F(t)$$

$$x(0) = 0, \quad \frac{dx}{dt}|_{x=0} = 0, \quad F(t) = 0 \quad t < 0$$

can be found by Fourier representing  $F(t)$ . Show that

$$x(t) = \frac{-1}{\sqrt{2\pi m}} \int_{-\infty}^{\infty} \frac{F(\omega)e^{i\omega t}}{\omega^2 - 2i\alpha\omega - \omega_0^2} d\omega ,$$

where  $F(\omega)$  is the FT of  $F(t)$ .

b) Use the Faltungstheorem, Problem 1.15, in conjunction with (2.28) and Part (a) here to show that the Fourier transform of  $G(t)$ ,  $G(\omega)$  is given by:

$$G(\omega) = \frac{-1}{\sqrt{2\pi} m(\omega^2 - 2i\alpha\omega - \omega_0^2)} . \quad (2.32)$$

c) Use whatever results in Sect. 2.5 you may need to show that:

$$\int_{-\infty}^{\infty} \frac{e^{i\omega t}}{\omega^2 - 2i\alpha\omega - \omega_0^2} d\omega = -(2\pi/\omega_f)e^{-\alpha t} \sin \omega_f t \quad t > 0$$

$$= 0 \quad t < 0 .$$

This integral can be done directly by using complex variable analysis, thereby providing a check on (2.28).

**2.10** The AC electric susceptibility  $\chi$  is defined by the equation

$$P(\omega) = \chi(\omega)\epsilon_0 E(\omega),$$

where  $P$ , the dipole moment per unit volume, in a simple model of polarization is given by

$$P(\omega) = N e x(\omega).$$

Here there are  $N$  ions per unit volume, each with a charge  $e$  oscillating harmonically about  $x(\omega) = 0$ .

Referring to Problem 2.9(d) show that

$$\chi(\omega) = (N e^2 / \epsilon_0) G(\omega),$$

Let  $\chi(\omega) = \chi_1(\omega) + i\chi_2(\omega)$  where  $\chi_1$  and  $\chi_2$  are real.

Find  $\chi_1$  and  $\chi_2$  as functions of  $\omega$ , and sketch the results. Which of the two functions,  $\chi_1(\omega)$  or  $\chi_2(\omega)$ , is associated with power loss?

**2.11** Find the complex impedance  $Z$  for an idealized circuit with a capacitor and an inductor in parallel. Start with the appropriate differential equations, or use any other method you may know to find  $Z$ . The solution of this problem will be useful later in the development of this text, Problems 3.15, 4.21.

## Further Reading

R. Resnick, D. Halliday: *Physics*, Vols. I and II (Wiley, New York 1977)  
An introduction to harmonic oscillators is given in these volumes.

I.G. Main: *Vibrations and Waves in Physics* (Cambridge University Press, Cambridge 1978)

A thorough treatment of all degrees of oscillator damping is found here.

P.R. Wallace: *Mathematical Analysis of Physical Problems* (Dover, New York 1984)

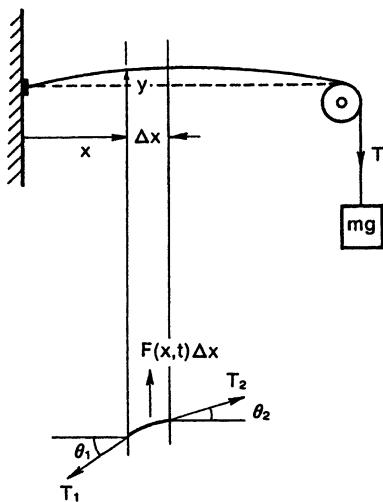
A full treatment of the oscillator in the content of Fourier analysis can be found in this text.

### 3. Waves on Stretched Strings

**Abstract.** The string under tension, when only lightly disturbed, is the simplest example of a distributed system undergoing linear wave motion. This chapter begins with a derivation of the wave equation governing the vibrations of the string. As in Chap. 2 we divide the analysis into natural and driven motion. The natural motion is shown to consist of an infinite number of independent contributions, each contribution a so-called normal mode. Associated with each mode is a harmonic oscillator, with a natural frequency characteristic of the mode. The spatial dependences of the modes together form a complete orthonormal basis over the interval occupied by the string. This last fact facilitates the solution of the initial value problem. Analysis of the driven motion reveals that the oscillator characterizing each mode displays a resonance response just as did the single oscillator in the previous chapter. The analysis is shown to be an instance of the systematic treatment of driven motion using Green's function techniques. The techniques are likewise applied to the problem of a stretched string driven from one end.

#### 3.1 Equation of Motion of a String

In the previous chapter we distinguished between lumped and distributed systems. The simplest example of a distributed system is one which extends in only one dimension, in mechanics we have the flexible string under tension. The string of a violin is an obvious application, human vocal chords are an application. We will study this problem in detail because it contains all the elements of analysis inherent in any distributed system with, however, a minimum of complications. As it is, the problems of this chapter are considerably more complex than those of the previous. The reason is not hard to find: If we were to divide the string into its component atoms we could study the motion of each atom by the methods of the previous chapter. But there would be very many atoms, very many lumped systems of which to keep track, very many independent variables or *degrees of freedom*. Instead of doing such an atomistic calculation one analyzes the motion of the string by considering it to be a single continuous system. Even so, characterizing the motion can involve the time dependence of an infinite set of variables, in correspondence to an infinite number of degrees of freedom.



**Fig. 3.1.** *Top:* The flexible string slightly displaced from equilibrium under a tension  $T$ . The pulley is frictionless. *Below:* The segment  $\Delta x$  of the string under a vertical force per unit length  $F(x, t)$

As in the previous chapter, Newton's second law will give us an equation of motion whose solutions will correctly predict the motion of the system under specified imposed conditions. In Fig. 3.1 we show the flexible string under tension  $T$  slightly displaced from its equilibrium configuration.

As shown, any point on the string can be identified by the value of the coordinate  $x$  which measures the horizontal distance from the beginning of the string. If the string has length  $L$ ,  $x$  goes from zero to  $L$ . The vertical or *transverse* displacement of the string at any point is measured by a second coordinate,  $y$ . The latter will depend on two variables, we write  $y(x, t)$ ,  $t$  being the time.

In contrast to the previous chapter we have to deal with functions of more than one variable. When we take the derivative of these functions it is the *partial* derivative that we usually need. The partial derivatives of  $y(x, t)$  are defined as

$$\frac{\partial y(x, t)}{\partial x} = \lim_{\Delta x \rightarrow 0} 1/\Delta x [y(x + \Delta x, t) - y(x, t)] , \quad (3.1a)$$

$$\frac{\partial y(x, t)}{\partial t} = \lim_{\Delta t \rightarrow 0} 1/\Delta t [y(x, t + \Delta t) - y(x, t)] . \quad (3.1b)$$

The partial derivative  $\partial y / \partial x$  is the usual derivative of  $y$  with respect to  $x$ , with the other variable,  $t$ , remaining a constant parameter. As we might expect, the equation of motion of the string will be in terms of partial derivatives, a partial differential equation. Indeed, distributed systems satisfy partial differential equations, and we must learn how to solve them.

To find the equation of the string we analyze the problem by isolating that segment of the string which stretches from  $x$  to  $x + \Delta x$ , Fig. 3.1. What will be the net force acting on this segment? We know that along a flexible string

the tension always acts in a direction tangent to the string. In the present problem it is important to distinguish between the slopes of the string at the two ends of the segment; these are, respectively, at some instant  $t$ ,

$$\tan \theta_2 = \frac{\partial y(x + \Delta x, t)}{\partial x} \quad \text{and} \quad \tan \theta_1 = \frac{\partial y(x, t)}{\partial x}.$$

We now make an important restriction which will hold throughout the chapter. We shall assume that we are concerned only with *small* vibrations of the string. The angles  $\theta_1$  and  $\theta_2$  will always be sufficiently small that in a Taylor expansion of the trigonometric functions we can stop at the first term. This restriction will ensure that our ensuing differential equation of motion is linear. We recall that this means that only the linear power of  $y(x, t)$  appears. Thus physically we are dealing, as so often is the case, with an *idealization*, an approximation, yet one that we expect will be very useful.

We see from Fig. 3.1 that the horizontal components of the forces acting on the string are, respectively,  $T_2 \cos \theta_2$  and  $T_1 \cos \theta_1$ . But, since  $\theta_1$  and  $\theta_2$  are to be small angles, we replace  $\cos \theta_1$  and  $\cos \theta_2$  by unity. Since the string does not move in the horizontal direction, we find

$$T_1 \cos \theta_1 = T_2 \cos \theta_2, T_1 = T_2 \equiv T.$$

The string has the same tension  $T$  all the way down its length, even though we are not going to neglect its mass.

The vertical components of the tension are  $T \sin \theta_1$ ,  $T \sin \theta_2$ . Again, since the angles are small, we can replace their sines by their tangents, so that the net vertical component of force exerted on the segment of the string by the two ends is

$$T(\sin \theta_2 - \sin \theta_1) = T \left[ \frac{\partial y(x + \Delta x, t)}{\partial x} - \frac{\partial y(x, t)}{\partial x} \right].$$

We shall be interested in the possibility of having some external force applied to the string. It will suffice for our purposes to restrict this force to act always in the vertical direction. If we say that a force  $F(x, t)$  *per unit length* is applied to the string at the point  $x$ , then the force acting on the segment is  $F(x, t)\Delta x$ .

We must also find an analytic expression for the mass acceleration term. We assume the string is uniform with a constant density, i.e., the mass per unit length is  $\varrho$ . The mass of the segment becomes  $\varrho\Delta x$ , the vertical acceleration of the segment is  $\partial^2 y(x, t)/\partial t^2$ , and Newton's second law applied to the vertical components gives us the desired equation of motion:

$$T \left[ \frac{\partial y(x + \Delta x)}{\partial x} - \frac{\partial y(x)}{\partial x} \right] + F(x, t)\Delta x = \varrho\Delta x \frac{\partial^2 y}{\partial t^2},$$

or

$$\frac{T}{\varrho} \frac{\partial^2 y(x, t)}{\partial x^2} + \frac{F(x, t)}{\varrho} = \frac{\partial^2 y(x, t)}{\partial t^2} \quad L > x > 0 \quad y \ll L. \quad (3.2)$$

In going from the first equation to (3.2) we have divided by  $\varrho\Delta x$ , gone to the limit  $\Delta x \rightarrow 0$ , and used (3.1a) for the result:

$$\lim_{\Delta x \rightarrow 0} 1/\Delta x \left[ \frac{\partial y(x + \Delta x, t)}{\partial x} - \frac{\partial y(x, t)}{\partial x} \right] = \frac{\partial^2 y(x, t)}{\partial x^2} .$$

## 3.2 Natural Motion of a String

As in the previous chapter we shall distinguish between two problems, the natural motion of the string, that is, its motion when after an initial disturbance the string is left to itself without any external applied force, and the forced motion of the string under the influence of a prescribed driving force  $F(x, t)$ . The natural or forced motion of any distributed system may be referred to as *wave motion*. Usually the first wave motion we observe as children is that of the surfaces of bodies of water. These are transverse waves, the physical displacements of the water being perpendicular to the direction of motion (*propagation*) of the waves. The waves in this chapter are likewise transverse and, being linear, are simpler to analyze. Surface water waves are included in the last chapter, where our work is extended to nonlinear waves.

Since we are not to have any external force in this section the problem is to understand the implications of the equation

$$\frac{\partial^2 y(x, t)}{\partial t^2} = s^2 \frac{\partial^2 y(x, t)}{\partial x^2} \quad (3.3)$$

obtained from (3.2) by setting  $F(x, t)$  equal to zero, and by making the standard substitution

$$s^2 = T/\varrho . \quad (3.4)$$

The partial differential equation (3.2) is a *wave equation*, indeed, it can probably be considered as the most basic of all the wave equations studied in physics.

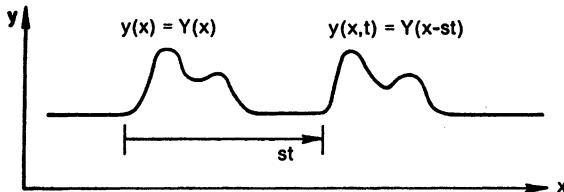
One way of writing the solution to this wave equation is surprisingly general in character:

$$y(x, t) = Y_1(x - st) + Y_2(x + st) . \quad (3.5)$$

Here  $Y_1$  and  $Y_2$  are any reasonable functions whatever. By “reasonable” we mean continuous enough to be assumed by a physical string without its having to be cut.<sup>1</sup> What does a solution  $Y(x - st)$  mean? Suppose at  $t = 0$  the string is deformed into a given arbitrary shape  $Y(x)$ . Then, at a later time  $t$ , the string will have exactly the same shape, but evidently a given feature is found at a point a distance  $st$  farther to the right. We illustrate this in Fig. 3.2. We

---

<sup>1</sup> The proof of (3.5) is simple. Let  $\xi = x - st$ . Then  $\partial^2 Y(x - st)/\partial t^2 = s^2 Y''$ , where  $Y''$  is the second derivative of  $Y(\xi)$  with respect to  $\xi$ . Similarly,  $\partial^2 Y(x - st)/\partial x^2 = Y''$ .

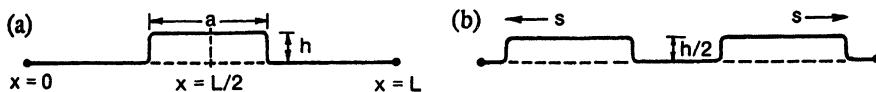


**Fig. 3.2.** A possible solution to the wave equation. Wave is moving to the right

can say that whatever the nature of the disturbance, it moves with a uniform velocity  $s$  to the right. The solution  $Y(x+st)$  has a similar implication, only now the disturbance is moving to the left.

If we refer back to (3.4) we see that  $s$ , which we have just learned is the speed of wave motion as supported by the string, increases with the tension on the string, although only as the root, and decreases with its density in the same way. Surely, this is what we might have expected from our own daily experience of all manners of elastic responses.

Let us use the knowledge about the solutions of the wave equation we have obtained so far to study a particular problem: In Fig. 3.3a a string is displaced an amount  $h$  between  $x = L/2 - a/2$  and  $x = L/2 + a/2$ , and not displaced at all everywhere else. In other words, we have deformed the string into a “square wave”. What will be the natural motion of the string if it is displaced in this fashion and then released from rest? Will the pattern move to the left, or will it move to the right? When seen from the center of the string the situation looks exactly the same towards the right as the left. Therefore, the square wave will be divided into two waves, each of displacement  $h/2$ , one of these waves will move to the right, the other to the left. Figure 3.3b shows the situation after a short time interval. In Problem 3.1 we verify this solution analytically, and include the possibility of an initial velocity.



**Fig. 3.3.** (a) Displacement of stretched string, as initially at rest. (b) Motion of string in consequence of initial symmetry

But what happens when the moving square waves hit the ends of the string? We are to think of both ends as rigidly fixed. In other words the *boundary conditions* on the string are

$$y(0, t) = 0, \quad y(L, t) = 0. \quad (3.6)$$

There is a method of graphical analysis (the method of images) which will enable us to handle the effects that the rigid end points provide, see

Example 3.1 at the end of the chapter. However, sometimes, particularly if one is interested in the long time behavior, one needs a more analytic account of the solutions of the wave equation than we have so far seen. In our later work we shall be concerned with other distributed systems, and again more complete knowledge will be called for.

### 3.3 The Normal Modes

We can solve the wave equation (3.3) in rather a particular way. Since the ultimate solution we obtain with this starting point will in fact be the most general solution, we need not concern ourselves with other possible ways of going about solving the equation. Let us substitute for  $y(x, t)$  a product of two functions, one of which depends only on  $x$  and the other only on  $t$ :

$$y(x, t) = u(x)v(t). \quad (3.7)$$

If we substitute (3.7) directly into (3.3), we find

$$u(x) \frac{d^2v(t)}{dt^2} = s^2 v(t) \frac{d^2u(x)}{dx^2}.$$

We next divide both sides by the product  $u(x)v(t)$ :

$$\frac{1}{v(t)} \frac{d^2v(t)}{dt^2} = \frac{s^2}{u(x)} \frac{d^2u(x)}{dx^2}. \quad (3.8)$$

The last equation has a remarkable feature. The left side depends on  $t$ , but not on  $x$ , the right side depends on  $x$  but not on  $t$ , yet the two sides are supposed to be equal for every value of  $x$  and  $t$ . This can only be true if, in fact, each side is equal to the same constant, i.e., a quantity which depends neither on  $x$  nor on  $t$ . Anticipating the next steps, we will call this constant  $-\omega^2$ . We have, then,

$$\frac{1}{v(t)} \frac{d^2v(t)}{dt^2} = -\omega^2, \quad (3.9a)$$

$$\frac{s^2}{u(x)} \frac{d^2u(x)}{dx^2} = -\omega^2. \quad (3.9b)$$

Let us take the first of these:

$$\frac{d^2v}{dt^2} = -\omega^2 v(t), \quad (3.10a)$$

This is the equation of an undamped harmonic oscillator, already studied in Chap. 2. Its solution is

$$v(t) = \sin \omega t, \quad \text{or} \quad \cos \omega t, \quad \text{or} \quad e^{i\omega t}, \quad (3.11a)$$

or any arbitrary combination of these. We again have a function periodic in time, and  $\omega$  is revealed as an angular frequency.

Likewise (3.9b) reads

$$\frac{d^2u(x)}{dx^2} + \frac{\omega^2}{s^2}u(x) = 0, \quad (3.10b)$$

with the solution

$$u(x) = \sin(\omega/s)x \quad \text{or} \quad \cos(\omega/s)x \quad \text{or} \quad e^{i(\omega/s)x}. \quad (3.11b)$$

Since these functions are spatially periodic we follow Chap. 1 in defining a wave vector  $k$  by

$$k = \omega/s. \quad (3.12)$$

Summarizing up to this point, a solution of the wave equation (3.3) for the stretched string will be a product of a function periodic in space,  $u$ , and one periodic in time,  $v$ . Besides (3.3) we must also satisfy the *boundary conditions* given by (3.6). Since these must hold throughout time, we must rely on  $u(x)$ . The necessary and sufficient conditions are

$$u(0) = 0 \quad u(L) = 0. \quad (3.13)$$

It follows that in order that  $u$  should vanish at  $x = 0$  we must restrict the possible solutions given in (3.11b) to the sine function. To satisfy the condition at  $x = L$  we find that we must put restrictions on the wave vector  $k$ , it becomes limited to certain particular values. We have

$$u(x) = \sin k_n x,$$

$$u(L) = 0,$$

$$k_n = n\pi/L \quad n = 1, 2, 3 \dots. \quad (3.14a)$$

In (3.14a)  $n$  can be any positive integer. We discard the negative integers, since they just give us back the same function with a minus sign, i.e., the same solutions again.

In view of the relationship (3.12) we must now also restrict the possible values of the angular frequency to

$$\omega_n = sk_n = n\pi s/L. \quad (3.14b)$$

The solution of our wave equation now reads

$$y_n(x, t) = (A_n \sin \omega_n t + B_n \cos \omega_n t) \sqrt{\frac{2}{L}} \sin \left( \frac{n\pi x}{L} \right). \quad (3.15)$$

Here  $A_n$  and  $B_n$  are arbitrary constants. Equation (3.15) will be a solution of the wave equation (3.3) regardless of what value we ascribe to them.  $\sqrt{2/L}$  is a “normalizing” factor. We will see that having it there makes things neater. Lastly, we notice that the choice of a negative real constant in (3.9a,b) was necessary to secure physical “boundary” conditions in space and time.

Since the wave equation, as already mentioned, is linear, its solution will not only be of the form (3.15) where  $n$  is any positive integer, but also the sum of any number of such solutions:

$$y(x, t) = \sum_{n=1}^{\infty} (A_n \sin \omega_n t + B_n \cos \omega_n t) \sqrt{\frac{2}{L}} \sin \left( \frac{n\pi x}{L} \right). \quad (3.16a)$$

It turns out that with the doubly infinite set of constants  $A_n, B_n$  at our disposal we can correctly describe any natural motion that a stretched string with fixed ends can be made to undergo. We have found the most general solution of our wave equation subject to the imposed boundary conditions at  $x = 0$  and  $x = L$ . The actual values of  $A_n$  and  $B_n$  will depend on the *initial conditions*, on the shape, and on the distribution of speeds over the length of the string with which we initiate the subsequently free motion. We have an extension of the simple harmonic solution we found for one oscillator in the previous chapter. Again, before we can solve a specific problem, we need to complete our formal work by writing along with (3.16a) a corresponding solution for the velocity distribution over the string:

$$\frac{\partial y(x, t)}{\partial t} = \sum_{n=1}^{\infty} \omega_n (A_n \cos \omega_n t - B_n \sin \omega_n t) \sqrt{\frac{2}{L}} \sin \left( \frac{n\pi x}{L} \right). \quad (3.16b)$$

We return to the problem of Fig. 3.3. From (3.16a,b) we have at time  $t = 0$ :

$$y(x, 0) = \sum_n B_n \sqrt{\frac{2}{L}} \sin \left( \frac{n\pi x}{L} \right), \quad (3.17a)$$

$$\left. \frac{\partial y(x, t)}{\partial t} \right|_{t=0} = 0 = \sum_n A_n \omega_n \sqrt{\frac{2}{L}} \sin \left( \frac{n\pi x}{L} \right). \quad (3.17b)$$

The left side of (3.17a) is the function of  $x$  shown in Fig. 3.3, that is, the initial displacement of the string. Since we assume that the string starts from rest all along its length, we have put the left side of (3.17a) equal to zero. What we see on the right side are expressions very reminiscent of the Fourier representations of arbitrary functions we studied in Chap. 1. The differences are ones of detail; in Chap. 1 we were concerned with functions described over the infinite range  $\infty > x > -\infty$ , while here the range is finite:

$$L > x > 0.$$

In correspondence, there the range of  $k$  was over all possible values, whereas here  $k$  is restricted to certain discrete values, enumerated by the positive integers. The integral of Chap. 1 has become a discrete sum. But the general ideas are the same, and the constants  $A_n$  and  $B_n$  will be determined in much the same way.

The first fact we establish is that the functions  $u_n$

$$u_n = \sqrt{\frac{2}{L}} \sin k_n x = \sqrt{\frac{2}{L}} \sin \left( \frac{n\pi x}{L} \right) \quad n = 1, 2, 3 \dots \quad (3.18)$$

again form an orthogonal set:

$$\int_0^L u_n(x)u_m(x)dx = 0 \quad \text{when } n \neq m . \quad (3.19a)$$

Further, in virtue of the factor  $\sqrt{2/L}$ , integration also shows us that

$$\int_0^L u_n(x)u_m(x)dx = 1 \quad \text{when } n = m . \quad (3.19b)$$

Equation (3.19b) allows us to say that the functions  $u_n$  are normalized. We have already said that (3.16a) is the most general solution to our problem. Equations (3.17) can be satisfied for any physical choice of the functions  $y(x, 0)$ ,  $\partial y(x, 0)/\partial t$ . These facts lead us to say, as in Chap. 1, that the functions

$$u_n(x), \quad n = 1, 2, 3 \dots$$

form a *complete orthonormal* set of *basis* functions.

We now use (3.19) to solve (3.17). We multiply both sides of (3.17a) by  $\sqrt{2/L} \sin(m\pi x/L)$  and integrate both sides over the defined range of  $x$ :

$$\sqrt{\frac{2}{L}} \int_0^L \sin\left(\frac{m\pi x}{L}\right) y(x, 0) dx = \sum_{n=1}^{\infty} B_n \left(\frac{2}{L}\right) \int_0^L \sin\left(\frac{m\pi x}{L}\right) \sin\left(\frac{n\pi x}{L}\right) dx .$$

If we compare the right side of this equation with (3.19a), we see that every term in the sum will be zero except the one for which  $n = m$ . Further, (3.19b) shows us that this particular term in the sum is just  $B_m$ . So turning the equation around, we have the desired formula for the coefficient  $B_m$ :

$$B_m = \sqrt{\frac{2}{L}} \int_0^L \sin\left(\frac{m\pi x}{L}\right) y(x, 0) dx = \int_0^L u_m(x) y(x, 0) dx . \quad (3.20a)$$

A similar procedure applied to (3.17b) leads to the formula

$$\begin{aligned} A_m &= \frac{1}{\omega_m} \sqrt{\frac{2}{L}} \int_0^L \sin\left(\frac{m\pi x}{L}\right) \frac{\partial y(x, 0)}{\partial t} dx \\ &= \frac{1}{\omega_m} \int_0^L u_m(x) \frac{\partial y(x, 0)}{\partial t} dx . \end{aligned} \quad (3.20b)$$

If now at last we apply (3.20) to the particular function  $y(x, 0)$  of Fig. 3.3, then we may write

$$B_m = h \sqrt{\frac{2}{L}} \int_{L/2 - a/2}^{L/2 + a/2} \sin\left(\frac{m\pi x}{L}\right) dx ,$$

since  $y(x, 0)$  is zero everywhere else. Doing the integration we find,

$$B_m = \frac{2h\sqrt{2L}}{m\pi} \sin \frac{m\pi}{2} \sin\left(\frac{am}{L} \cdot \frac{\pi}{2}\right) . \quad (3.21a)$$

Since the initial velocity is zero everywhere, we have also, from (3.20b)

$$A_m = 0 \quad (3.21b)$$

for all  $m$ . We notice that  $B_m$  is also equal to zero for all even values of  $m$  in view of the factor  $\sin(m\pi/2)$  in (3.21a). The situation is reminiscent of that encountered in Example 1.1. The even symmetry about the center of the string excludes all the functions  $u_n$  that themselves have odd symmetry about the center.

Equations (3.16a, 21) constitute the solution to our *initial value* problem. When the coefficients  $A_m, B_m$  as obtained from (3.21) are substituted into the general expression (3.16a) we have the development of the displacement  $y$  at all points  $x$  throughout time. We see that given the problem of some distributed system such as our stretched string, with the positions and velocities specified at some initial time  $t = 0 \dots$  there is a systematic procedure for finding the motion of the system at all subsequent times. On physical grounds we expect that this solution is unique; once the system is started it always develops in the same way. Mathematics confirms that a partial differential equation of the nature of our wave equation, with given boundary conditions, has only one solution if the initial values of position and velocity have been specified. Since the preceding development has been a rather long one, we summarize it in Table 3.1.

There is one feature of (3.16a) which requires further comment. We notice that once the magnitude of the amplitudes  $A_n$  and  $B_n$  have been determined they remain constant throughout time. There is no cross-linkage between the various component solutions  $y_n$  of (3.15):

$$\begin{aligned} y_n(x, t) &= (A_n \sin \omega_n t + B_n \cos \omega_n t) u_n(x) , \\ y &= \sum_{n=1}^{\infty} y_n . \end{aligned} \quad (3.16a)$$

A quantity which shows this independence of the various components of the motion particularly clearly is the energy of the string, which shows up as a sum of terms, one for each value of  $n$ . This fact is taken up in one of the problems at the end of the chapter. In more advanced treatments it is possible to solve for the motion of systems by looking at the kinetic and potential energies rather than at forces. The problem then becomes one of determining functions  $y_n(x, t)$  which will lead to contributions independent of each other in the expression for the energy.

**Table 3.1.** Systematic procedure for solving an initial value problem for a distributed system

A string under tension  $T$  extends from  $x = 0$  to  $x = L$ . Its density is  $\varrho$ . The small displacement of the string,  $y$ , satisfies the wave equation

$$\frac{\partial^2 y(x, t)}{\partial t^2} = s^2 \frac{\partial^2 y(x, t)}{\partial x^2} \quad (3.3)$$

$$\text{with } s^2 = T/\varrho. \quad (3.4)$$

The boundary conditions are

$$y(0, t) = 0 \quad y(L, t) = 0. \quad (3.6)$$

The general solution satisfying (3.3) and (3.6) throughout time is

$$y(x, t) = \sum_{n=1}^{\infty} (A_n \sin \omega_n t + B_n \cos \omega_n t) u_n(x). \quad (3.16a)$$

Here  $\omega_n = sk_n$ ,  $u_n(x)$  is the complete, orthonormal set of functions

$$u_n(x) = \sqrt{\frac{2}{L}} \sin k_n x \quad (3.18)$$

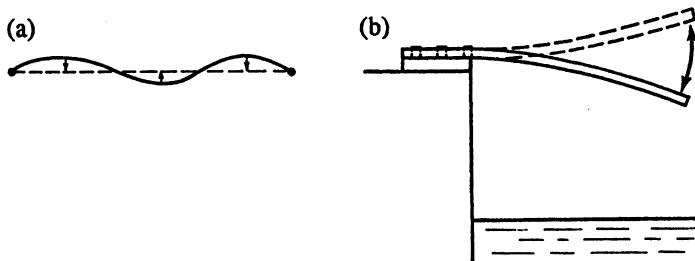
$$k_n = \frac{n\pi}{L}. \quad (3.14a)$$

The coefficients  $A_n, B_n$  are determined from the initial conditions through the formulas

$$A_n = \frac{1}{\omega_n} \int_0^L u_n(x) \frac{\partial y(x, t)}{\partial t} |_{t=0} dx, \quad (3.20a)$$

$$B_n = \int_0^L u_n(x) y(x, 0) dx. \quad (3.20b)$$

With all these properties, it is hardly surprising that we give the independent solutions of the problem,  $y_n$ , a name. They are the *normal modes* of motion of the system. A corollary of the solution (3.16a) is that if we were to excite the string so that only one mode, say  $y_1$ , were initially present, then the string will vibrate throughout time in only this one mode, Fig. 3.4. Often, in practice the higher modes damp out rather quickly, leaving only the lowest (fundamental) mode, see Fig. 3.4b. We say the modes are *characteristic* of the system. [In more advanced treatments one also uses the designation “eigenfunction” for the functions  $y_n(x, t)$ , *eigen* meaning “characteristic” in German. We will stress the concepts that go with the latter designation in Chap. 6, when we come to quantum mechanics.] We summarize by saying



**Fig. 3.4.** (a) Stretched string oscillating in the  $n = 3$  mode only. (b) Diving board oscillating in its lowest characteristic mode. Any higher modes initially excited would be damped out very rapidly

that the motion of distributed systems can be viewed as a superposition of basic modes of motion, the normal modes.

Let us examine the normal modes of the string in more detail. First, we notice that for all of them the time dependence is simple harmonic. In fact, the angular frequencies  $\omega_n$  are integral multiples of the lowest frequency  $\omega_1$ , which is called the *fundamental frequency* of the system. We have

$$\omega_n = n\omega_1 \quad n = 1, 2, 3, \dots \quad \omega_1 = \pi s/L .$$

The higher frequencies are sometimes referred to as the higher harmonics or overtones.

The normal modes we have found are examples of *standing waves*. We have seen that the definition of a wave is any disturbance (motion) supported by a given distributed system. For the homogeneous (constant density) string we have found that the normal modes are simple harmonic in both space and time. In elementary treatments waves are introduced from the beginning as having the functional form

$$y = A \begin{Bmatrix} \cos \omega t \\ \sin \omega t \end{Bmatrix} \begin{Bmatrix} \cos kx \\ \sin kx \end{Bmatrix} ,$$

i.e., as one of these four products. These waves are called *standing* because the positions of the crests and troughs of the waves are always at the same location in space.

There is one further relation characteristic of harmonic waves. Let  $f_n (= \omega_n/2\pi)$  be a characteristic frequency of a wave. Let  $\lambda_n$  [already introduced in (1.10) as  $2\pi/k_n$ ] be the *wavelength* associated with the wave. Then a little algebra shows us that (3.14b) can be written

$$\lambda_n f_n = s = \text{speed of wave} , \quad (3.22)$$

a result independent of  $n$ . Besides standing waves one can also talk about *traveling waves*,

$$y = A \sin(kx - \omega t) .$$

We note that these are a special example of the solution (3.5) of the wave equation. Here the relation  $\lambda f = s$  has an obvious physical interpretation. Further, two harmonic travelling waves of the same frequency but travelling in opposite directions lead to a standing wave. Since not every reader is assumed to be familiar with these basic ideas, we take them up again in Example 3.1 and in Problem 3.15.

### 3.4 Forced Motion of a Stretched String

Having solved the problem of the natural motion of the string, we now inquire into forced motion; how does the system respond to the action of a driving force? Since  $F(x, t)$  in (3.2) is now expected to be the controlling influence, we rewrite (3.2), recalling (3.4), as

$$\frac{\partial^2 y}{\partial t^2} + \beta \frac{\partial y}{\partial t} - s^2 \frac{\partial^2 y}{\partial x^2} = \frac{F(x, t)}{\varrho}. \quad (3.23)$$

Here we have added a damping force per unit length  $-(\beta \varrho \partial y / \partial t)$ , proportional to the velocity  $\partial y / \partial t$ , just as we added the damping force  $-bdx/dt$  to the harmonic oscillator in Chap. 1. However, we shall see that there will not be any need to repeat a full description of its effects. Rather, it is introduced for mathematical definition, i.e., to avoid again certain infinities, and will be taken as infinitesimally small in most of what follows.

We need not solve the problem of the string under an arbitrary force. Rather we will specialize to a localized force in space, and one which is simple harmonic in time. We can write a force localized in space analytically by the use of the Dirac delta function discussed in Chap. 1, and defined in (1.23). Accordingly, we have for the force per unit length  $F(x, t)$ ,

$$F(x, t) = F_\omega \delta(x - x_0) \cos \omega t. \quad (3.24)$$

Incidentally, the total force acting on the string now will be given by the integral of  $F(x, t)$  over the string by

$$\int_0^L F(x, t) dx = \int_0^L F_\omega \delta(x - x_0) \cos \omega t dx = F_\omega \cos \omega t.$$

If we combine (3.23) and (3.24) we have

$$\frac{\partial^2 y}{\partial t^2} + \beta \frac{\partial y}{\partial t} - s^2 \frac{\partial^2 y}{\partial x^2} = \frac{F_\omega}{\varrho} \delta(x - x_0) \cos \omega t. \quad (3.25)$$

If we want to know the response of the string to an arbitrary force in space and time, then we can always decompose that force into a sum of terms, each of the form of (3.24) but having various frequencies, and various locations  $x_0$  along the length of the string. We can then use a feature of inhomogeneous linear equations already pointed out in the previous chapter, namely, that the

solution of such equations is obtained by adding together individual solutions, each of which is obtained by taking the right side of the equation to be a particular one of the component driving terms. We shall not go so far here, leaving that to the Problems, but rather try to understand well the effect of a force of the form (3.24).

What will be the response in time of a system under a simple harmonic driving force, as in (3.25)? First, we must distinguish between the steady state response and the startup problem. The latter introduces that solution to the homogeneous wave equation which must finally be added in to yield correspondence with the actual starting conditions. In the analysis of this start up, the initial value problem of the previous sections, we found that all the various normal modes of motion are excited to varying degrees, and the system will oscillate in all these modes. However, in that analysis we did not include frictional damping. The damping action will qualitatively be the same as it was for the single oscillator in Chap. 2. In practice, first the higher, faster modes of motion dissipate their energy, to be followed finally by the fundamental mode itself. We again refer to these effects as the “transient”, and find that their lifetimes, as for the single oscillator, is related reciprocally to the strength of the frictional damping  $\beta$ . What is left when the transient has died out? The forced motion, that is, that motion which depends solely on the steady, continuous harmonic driving force. As far as this steady state is concerned, we are to think of the situation as though the driving force has been present for all time, since negative infinity.

If the driving force has really been there since negative infinity, and will always be there, then we know what the time response of the system will be. We recall from Chap. 1 that functions of the form  $\cos \omega t$ ,  $\exp(i\omega t)$  are orthogonal over the infinite time domain. A necessary corollary is that the time response of  $y$  will likewise be simple harmonic with the same frequency as the driving force, exactly as for the driven oscillator of Chap. 2. If we neglect any phase shift due to damping,  $\beta$  infinitesimal only, then we will have  $y(x, t) \propto \cos \omega t$ . For the moment, more generally,

$$y(x, t) \propto e^{i\omega t}. \quad (3.26)$$

What about the spatial dependence of  $y$  on  $x$  for the forced motion? Here we can return to the methods of the previous sections, and recollect that the set of functions  $u_n(x)$  of (3.18) form a *complete* set of functions for the interval  $x = 0$  to  $x = L$ , providing that the end points are held fixed. That being the case, and bearing (3.26) in mind, we write  $y(x, t)$  in the form

$$y(x, t) = \sum_{n=1}^{\infty} C_n u_n(x) e^{i\omega t}. \quad (3.27)$$

As in previous sections the  $C_n$  are a set of (complex) coefficients that determine the sought-for solution [given by the real part of (3.27)].

Let us substitute (3.27) into (3.25):

$$\sum_{n=1}^{\infty} C_n \left[ u_n(x) \left( \frac{d^2}{dt^2} + \beta \frac{d}{dt} \right) - s^2 \frac{d^2 u_n(x)}{dx^2} \right] e^{i\omega t} = \frac{F_\omega}{\varrho} \delta(x - x_0) e^{i\omega t}.$$

Upon performing the necessary differentiation we find

$$-\sum_{n=1}^{\infty} C_n u_n(x) (\omega^2 - i\omega\beta) e^{i\omega t} + s^2 \sum_{n=1}^{\infty} C_n k_n^2 u_n(x) e^{i\omega t} = \frac{F_\omega}{\varrho} \delta(x - x_0) e^{i\omega t}.$$

If we cancel the common factor  $\exp(i\omega t)$ , the equation becomes

$$\sum_n C_n u_n(x) (s^2 k_n^2 + i\omega\beta - \omega^2) = \frac{F_\omega}{\varrho} \delta(x - x_0).$$

We now use the usual stratagem of multiplying both sides by  $u_m(x)$  and integrating over  $x$  from 0 to  $L$ :

$$\sum_n C_n (s^2 k_n^2 + i\omega\beta - \omega^2) \int_0^L u_m(x) u_n(x) dx = \frac{F_\omega}{\varrho} \int_0^L u_m(x) \delta(x - x_0) dx.$$

The left side reduces to a single term in view of the orthonormality conditions (3.19a,b); the right side is evaluated using the definition of the Dirac delta function, (1.23). We get

$$C_m (s^2 k_m^2 + i\omega\beta - \omega^2) = \frac{F_\omega}{\varrho} u_m(x_0).$$

Referring back to (3.14b), namely  $\omega_m = sk_m$ , we write  $C_m$  as

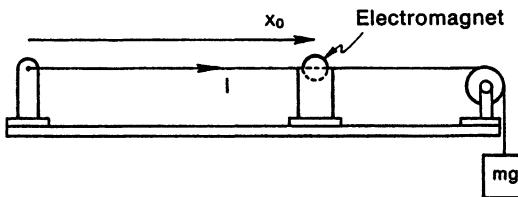
$$C_m = \frac{(F_\omega/\varrho) u_m(x_0)}{(\omega_m^2 + i\omega\beta - \omega^2)}.$$

When we substitute this result for  $C_m$  back into (3.27) and take the real value we get the desired response of the string. In Chap. 2 we have explored the full effect of damping,  $\beta \neq 0$ . Here  $\beta$  will be taken to be very small, so that if we are not exactly at a resonance we may, taking the real part, write for  $y(x, t)$

$$y(x, t) = \frac{F_\omega}{\varrho} \sum_n \frac{u_n(x) u_n(x_0)}{(\omega_n^2 - \omega^2)} \cos \omega t \quad (3.28)$$

$\omega \neq \omega_n, n = 1, 2, 3 \dots$

An apparatus to test (3.28) can be easily constructed. Figure 3.5 shows a schematic diagram. A light wire has its left end fixed. The other end of the wire is placed over a pulley at the far side of the platform, and has a hook for holding various weights attached to it. The apparatus is also provided with a small electromagnet, which itself produces an oscillating magnetic field, say at the line frequency of  $(2\pi)60$  rad/s. A DC electric current is sent down



**Fig. 3.5.** The wire carries a DC current  $I$ . The Lorentz force,  $Il \times B$ , exerted on the wire at  $x_0$  drives it vertically at the line frequency. The apparatus allows verification of the analysis of the forced motion of a stretched string

the wire, so that a Lorentz force ( $Il \times B$ ) will drive the wire at the line frequency in a direction perpendicular to the wire and the applied magnetic field. We identify the location of the magnet, which is adjustable, with  $x_0$ , and the force per unit length of the magnet with  $\delta(x - x_0)F_\omega \cos \omega t$  (total force  $F_\omega \cos \omega t$ ). The delta function representation of the exerted force is valid providing the magnetic field extends over a length small compared to the excited wavelengths of interest.

No matter where we apply the magnet, or what the tension on the wire, there will be some oscillation, and a low hum will be heard. We can always check out that the oscillating frequency is that in the line with a stroboscope, thereby verifying the last factor,  $\cos \omega t$ , in (3.28), but it would be surprising if the frequency of oscillation of the wire were anything different! Of course higher harmonics in the line voltage itself are always a possibility.

Now if we begin to adjust the tension, the amplitude of the wire oscillations will suddenly begin to increase drastically. This means that we are going through one of the nulls in the denominator in (3.28), that is, we have adjusted the tension so that

$$n \left( \frac{\pi}{L} \right) \sqrt{\frac{T}{\rho}} = \omega_n = \omega. \quad (3.29)$$

The system has reached a condition of *resonance*. The situation here is just that already encountered for the lumped single oscillator, only here we have an infinite number of resonances, one for each normal mode  $n$ . Had we included damping explicitly, as we did in the previous chapter, the denominators in (3.28) would have the more complicated form of (2.24c), and would remain finite at resonance.

Besides calculating  $n$  from (3.29), and verifying that it is an integer, with careful observation, or with the help of little paper riders on the wire, we can establish that the oscillations of the wire are a standing wave pattern described exactly by the corresponding function  $u_n(x)$  appearing in (3.28). From (3.28) we can also understand another observation, that is, that the response of the wire is greatest when the magnet is held at the precise location of one of the crests of the standing wave oscillation. This is predicted by the presence of the other factor,  $u_n(x_0)$ , in the numerator. Evidently that factor

determines the effectiveness of the magnet to system coupling. One does not usually trouble oneself to verify the presence of the first factor  $F_\omega/\rho$  in (3.28). But certainly its presence seems reasonable on physical grounds. Thus our mathematical reasoning can be tested out easily and completely, and is seen to provide an understanding of the entire phenomenon.

### 3.5 Green's Functions for a Stretched String

The calculation of the forced motion of a string carried out in Sect. 3.4 can be viewed within the broader theory concerned with the application of Green's functions. In the theory the Green's function technique for solving forced motion problems is extended from lumped systems dependent only on the time variable  $t$ , as in Chap. 2, to distributed systems dependent, in three dimensions, on a space variable  $\mathbf{r}$  as well as on  $t$ . The basic idea is as before. To solve the problem of a driven system, we first solve the problem with a drive that has a delta function distribution. The solution of this preliminary problem is the Green's function. The influence of boundaries aside, linear superposition then yields the solution to the full problem. In the last chapter the delta distribution was in time, here it could be in space, or time, or both. In fact, the distribution will be in space only, the treatment being restricted to a harmonic time dependence, with angular frequency  $\omega$ . Explicit subscripts referring to  $\omega$  will be omitted. We shall, accordingly, consider Green's functions  $G(\mathbf{r}|\mathbf{r}_0)$ , where  $\mathbf{r}$  is the location of the observer, and  $\mathbf{r}_0$  the location of the delta function source, i.e. of the driving force in mechanics, or of electrical charges or currents in electromagnetism.

The desired solution is going to depend not only on the actual drive, but on the boundary conditions of the system. For instance, if a stretched string has driven end supports, this will clearly affect the motion. The Green's function  $G(\mathbf{r}|\mathbf{r}_0)$  satisfies, by definition, so called *homogeneous* boundary conditions. Homogeneous boundary conditions means that when  $\mathbf{r}_0$  is at the boundary, i.e. at a point on the boundary  $\mathbf{r}_s$ , either  $G(\mathbf{r}|\mathbf{r}_s)$ , or the normal derivative of  $G$  with respect to  $\mathbf{r}_0$  at  $\mathbf{r}_s$ , (or possibly some linear combination of both these) vanishes. However, the *problems* to be solved, *may or may not* satisfy homogeneous boundary conditions (a driven end point would not be a homogeneous boundary condition). To see how, nevertheless, solutions are possible, we must derive a general relationship mediated by the Green's function between the desired solution,  $u(\mathbf{r})$ , its sources within some specified boundary, and the values of  $u(\mathbf{r}_0)$  and its derivative at the boundary.

In this chapter we shall restrict ourselves to one spatial dimension,  $x$ . The generalization of the relationship we are about to derive to three dimensions will be given in Chap. 5, where we shall see that it forms the analytic basis for doing physical optics. Working here in one dimension on the problem of stretched strings will make it easier to assess the three dimensional results when we reach Chap. 5.

As a first step we derive a well known formula, using standard integration by parts (chain rule). Let  $U(x)$ ,  $V(x)$  be any good functions of  $x$  in the range,  $L > x > 0$ . Then:

$$\int_0^L U(x) \frac{d^2 V(x)}{dx^2} - \int_0^L V(x) \frac{d^2 U(x)}{dx^2} dx = \left[ U(x) \frac{dV(x)}{dx} - V(x) \frac{dU(x)}{dx} \right]_0^L \quad (3.30)$$

after cancelling the extra term. (3.30) is the one dimensional equivalent of Green's theorem in three dimensional vector calculus. Now, let  $\tilde{y}(x, t|x_0)$  be the displacement of the string when it is driven harmonically from the point  $x_0$ , as in the last section, and  $y(x, t)$  the solution with an arbitrary harmonic driving force. Thus,  $\tilde{y}$  satisfies the equation of motion (3.25) with only the  $\beta$  (frictional) term left out, and will be given by:

$$\frac{\partial^2 \tilde{y}}{\partial t^2} - s^2 \frac{\partial^2 \tilde{y}}{\partial x^2} = F_\omega / \varrho \delta(x - x_0) \cos \omega t \quad (3.31)$$

$$\omega = sk \quad (3.12)$$

To obtain a conventional Green's function  $G(x|x_0)$  we write:

$$G(x|x_0) \cos \omega t \equiv (\varrho s^2 / F_\omega) \tilde{y}(x, t|x_0) \quad (3.32)$$

From (3.31) and (3.32) we obtain:

$$\frac{\partial^2 G(x|x_0)}{\partial x^2} + k^2 G(x|x_0) = -\delta(x - x_0) \quad (3.33)$$

(3.33) is the wave equation with a point source, which along with boundary conditions defines the Green's function. The specifics of the homogeneous boundary conditions of the Green's function will depend on the nature of the problem to be solved, as we shall see shortly. Again, following convention, we choose a formal solution  $u(x)$  proportional to the actual displacement of the string  $y(x, t)$ , where, again,  $y(x, t)$  is assumed to arise from a time harmonic drive:

$$u(x) \cos \omega t \equiv s^2 y(x, t) \quad (3.34)$$

$y(x, t)$  is to be the solution of the general equation of motion (3.23) with driving force  $F(x, t)$ , and the frictional coefficient,  $\beta = 0$ . Noting that the applied force is to be harmonic we have:

$$F(x, t) = f(x) \cos \omega t \quad (3.35)$$

Substituting (3.34) and (3.35) into (3.23) we obtain for  $u(x)$ :

$$\frac{d^2 u(x)}{dx^2} + k^2 u(x) = -f(x) / \varrho \quad (3.36)$$

It is (3.36), subject to boundary conditions on  $u(x)$  at  $x = 0, x = L$ , that we wish to solve. Equations (3.33) for the Green's function, and (3.30), Green's theorem are to help us.

Green's functions satisfy a reciprocity relation. In one dimension:

$$G(x|x_0) = G(x_0|x) \quad (3.37a)$$

We will take this relation for granted. It may be taken as a generalization of Newton's third law, "to every action there is an equal and opposite reaction". The reader who wants to see a mathematical proof (it is not long) is referred to the book by Morse and Feshbach listed at the end of the chapter. In any case we can always check that the Green's functions that are used here satisfy (3.37a). We shall also need the fact that:

$$\frac{\partial^2 G(x|x_0)}{\partial x^2} = \frac{\partial^2 G(x|x_0)}{\partial x_0^2} \quad (3.37b)$$

(3.37b) follows from (3.37a) and the definition of the Green's function (3.33). [Formally interchange  $x$  and  $x_0$  in (3.33), then use the reciprocity relation, and note that  $\partial(x - x_0) = \partial(x_0 - x)$ ].

If we now replace  $x$  by  $x_0$  in (3.36), and multiply the result through by  $G(x|x_0)$  we obtain:

$$G(x|x_0) \frac{\partial^2 u(x_0)}{\partial x_0^2} + G(x|x_0) k^2 u(x_0) = -G(x|x_0) f(x_0)/\varrho \quad (3.38a)$$

Now we multiply the equation for the Green's function (3.33) by  $u(x_0)$ :

$$u(x_0) \frac{\partial^2 G(x|x_0)}{\partial x^2} + G(x|x_0) k^2 u(x_0) = -u(x_0) \delta(x - x_0) \quad (3.38b)$$

Subtracting (3.38b) from (3.38a), and applying (3.37b) to the relevant term we obtain:

$$G(x|x_0) \frac{\partial^2 u(x_0)}{\partial x_0^2} - u(x_0) \frac{\partial^2 G(x|x_0)}{\partial x_0^2} = u(x_0) \delta(x - x_0) - f(x_0) G(x|x_0)/\varrho \quad (3.39)$$

Finally, we integrate both sides of (3.39) over the range,  $L > x > 0$ , and use Green's theorem (3.30):

$$\left[ G(x|x_0) \frac{\partial u(x_0)}{\partial x_0} - u(x_0) \frac{\partial G(x|x_0)}{\partial x_0} \right]_{x_0=0}^{x_0=L} + 1/\varrho \int_0^L G(x|x_0) f(x_0) dx_0 = u(x) \quad L > x > 0 \quad (3.40)$$

(3.40) is so useful, it is sometimes called the *magic rule*.

To apply the magic rule the boundary conditions of  $G(x|x_0)$  need to be chosen according to the conditions on  $u(x)$ , that is, at  $x = 0$  and  $x = L$ , (or, if one is working in three dimensions, on  $u(\mathbf{r})$  over a closed surface or at infinity). There are two possibilities in a given problem, either  $u$ , or the gradient of  $u$  is specified at the boundary. There is also the possibility that some mixture of both  $u$  and its gradient is given. However, this more general

case goes beyond what we need here or in subsequent chapters. Let us now suppose that it is  $u(x)$  that is specified at an end point. Then we must choose  $G(x|x_0)$  to vanish at that end point. As we are learning, this will (along with the condition at the other end) uniquely determine  $G$ , and, hence, uniquely  $u(x)$  as well. Here  $\partial u(x)/\partial x$  is initially unknown, and must be left free to be determined by the final solution for  $u(x)$ . Similarly, if it is  $\partial u(x)/\partial x$  that is specified at an end point, we choose  $G$  so that  $\partial G(x|x_0)/\partial x_0$  ( $x_0$  = end point) vanishes. It being both a mathematical and a physical fact that the motion of the string is fully determined by only one boundary condition, i.e. either value or slope at the end points, it follows that our specific choices of Green's functions have not introduced any lack of generality.

To make the procedure of solving problems with inhomogeneous boundary conditions using Green's functions, that satisfy, by definition, homogeneous boundary conditions convincing, we discuss the physical meaning of the terms on the left side of (3.40), one at a time. Beginning with the last term on the left, we remember that  $G(x|x_0)$  is the solution of the wave equation for the delta function source  $\delta(x - x_0)$ . The term, which can be seen to express summation or integration over the distributed source  $f(x_0)/\varrho$  follows from linear superposition.

To get at the second last term we consider the problem where at, say,  $x_0 = 0$  the displacement,  $u$  does not vanish, i.e. the string is driven from an end point. Drawing on the suggested procedure for this boundary condition, we choose  $G(x|0)$  to vanish. To simulate the drive at  $x = 0$ , we apply a strong delta function force at  $x_0 = \Delta$ ,  $\Delta \ll L$ , while assuming that  $u(0) = 0$ . This force is to give the correct  $u(x)$  for all  $x > \Delta$ . Now, see Problem (3.14b), a transverse force on the string,  $T(\tan \theta)$ , i.e.  $Ty/\Delta$ , or  $Tu(\Delta)/(s^2 \Delta) \cos \omega t$ , see (3.34) will provide the desired displacement  $u(\Delta)/s^2$ .  $T$  is the tension in the string. To obtain the source strength we must still divide the hypothetical force by  $\varrho$ , see (3.36). We are, thereby, left with a source strength  $u(0)/\Delta$ . If this source is multiplied by  $G(x|\Delta)$ , we can expect that the product will yield the actual displacement  $u(x)$  for all  $x > \Delta$ . Recalling that  $G(x|0)$  vanishes, we estimate  $G(x|\Delta)$  by  $\partial G(x|x_0)/\partial x_0 \Delta$ , at  $x_0 = 0$ . Finally, we multiply this estimate for  $G$  by the source strength  $u(0)/\Delta$ , and then go to the limit  $\Delta \rightarrow 0$ . In this way we recapture the second term on the left of the magic rule (3.40) when  $x_0 = 0$ .

Somewhat similar plausibility arguments can be used to explain the first term in the magic rule (3.40),  $G(x|x_0)\partial u(x_0)/\partial x_0$ , to be evaluated at the end points, see Problem (3.18). The preceding discussion interpreting the terms in the magic rule will be more readily convincing for readers familiar with calculating electrostatic potentials in three dimensions. The analogue of the simulating force is a charged surface lying just in front of the boundary surface. In general, then, boundary conditions can be replaced by equivalent sources at the boundaries. Indeed, turning things around, boundary conditions dictated by the nature of the physical apparatus are, in fact, brought

about by appropriate source strengths lying just in the physical boundaries. This fact will be very evident in the next chapter on electromagnetism, and can be taken as the physical origin of the “magic”.

It will be best to begin illustrating the use of the magic rule by considering the problem we have already solved in Sect. 3.3. Accordingly, we consider the stretched string fixed at  $x = 0$  and at  $x = L$ , driven by a transverse harmonic force with a delta function distribution at  $x = x_0$ . Referring to the right side of (3.28) in the previous section, actually the result for  $\tilde{y}$  when the end points are fixed, and (2.32) which defines  $G(x|x_0)$ , we find immediately that:

$$G(x|x_0) = \sum_{n=1}^{\infty} \frac{u_n(x)u_n(x_0)}{k_n^2 - k^2} \quad k \neq k_n \quad (3.41)$$

$$u_n(x) = \sqrt{2/L} \sin k_n x \quad (3.18)$$

$$k_n = n\pi/L \quad n = 1, 2, 3, \dots \quad (3.14a)$$

The solution (3.41) satisfies the wave equation, either (3.25) with  $\beta = 0$ , or, equivalently, (3.33). It also satisfies the homogeneous boundary conditions, with  $G(0|x)$ ,  $G(L|x)$  both vanishing. We see that the expression (3.28), derived in the previous section yields precisely the Green's function as defined by (3.33) and by the boundary conditions appropriate for the present problem. The magic rule now gives as the solution with the general force,  $f(x)$ :

$$u(x) = 1/\varrho \int_0^L G(x|x_0)f(x_0)dx_0 \quad L > x > 0 \quad (3.42a)$$

$$= 0 \quad x > L, x < 0 \quad (3.42b)$$

The boundary conditions on  $u$ , and on  $G(x|x_0)$  at  $x = 0$  and  $x = L$ , i.e. both quantities vanish there, have led to the vanishing of the boundary term on the left of (3.40).

As a second illustration we consider the situation in Problems 3.15, 3.16, namely a stretched string driven from the  $x = 0$  end:  $y(0,t) = A \cos \omega t$ . The solution can be obtained directly by the method of images, as in Problem (3.15). In Problem (3.16) we obtain the solution in terms of expansion in normal modes using a trick. Here we show how to proceed systematically using the magic rule.

From (3.34), which relates  $y(x,t)$  to  $u(x)$  we have that at the beginning support:

$$u(x_0) = As^2, \quad x_0 = 0 \quad (3.43a)$$

With the  $x = L$  point fixed we have

$$u(x_0) = 0, \quad x_0 = L \quad (3.43b)$$

Since the boundary conditions specify the values of  $u(x)$ , rather than its gradient, we are to choose the Green's function for which  $G(x|0)$  and  $G(x|L)$  both vanish. But this is just the Green's function we have in (3.41a,b). From the magic rule we obtain:

$$u(x) = u(0) \left. \frac{\partial G(x|x_0)}{\partial x_0} \right|_{x_0=0} \quad (3.44)$$

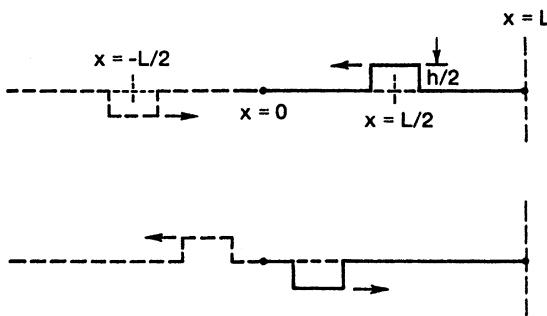
(3.44) reflects the fact that the only forces on the string are at the end points. Finally, combining (3.41) for  $G$  with (3.34) relating  $y$  to  $u$ , and remembering (3.43a) we find:

$$y(x, t) = \sqrt{\frac{2}{L}} A \sum_{n=1}^{\infty} \frac{k_n u_n(x)}{k_n^2 - k^2} \cos \omega t \quad L > x > 0 \quad (3.45)$$

This result agrees with that obtained in Problem (3.16), to which the reader has already been referred.

As  $x \rightarrow 0$ , the expression (3.45) does not converge. The origin of the difficulty is the fact that  $G(x|x_0)$  can not be differentiated at  $x = x_0$ . The situation is reminiscent of convergence difficulties when Fourier transforming a square wave, as in Example 1.2. There it was found that the Gibbs' phenomenon could be avoided by first smoothing the abrupt edges of the square wave (converting the square wave to a good function, Problem 1.18). Here we could, in principle, replace  $\delta(x - x_0)$ , the drive on the right side of the equation (3.33) for  $G(x|x_0)$ , by  $g_m(x - x_0)$ ,  $g_m(x)$  defined in (1.20) as  $\sqrt{m/\pi} \exp(-mx^2)$ . We should, then, get converging series in (3.45) even at  $x = x_0$ . The resulting sum for  $y(x, t)$  would be adequately represented by including wavelengths  $\lambda_n$  up to  $\lambda_n^2 < 1/m$ , where  $\lambda_n = 2\pi/k_n$  in (3.45). Thus, we must have  $k_n/(2\pi) > \sqrt{m}$ , or  $n(m) > L\sqrt{m}$ . To get results in (3.45) independent of  $m$  we would have to choose  $m$  larger and larger as we edge  $x$  closer to  $x_0$  ( $x_0$  being, say, the location of a simulating drive at  $x_0 \approx 0$ ). For a given difference  $|x - x_0| \geq \Delta$ ,  $m$  must exceed  $1/\Delta^2$ . This, finally, gives  $n \gg L/\Delta$ . The argument suggests that we shall get realistic results from (3.45) for  $y(x, t)$  as it stands providing we take terms up to  $n \gg L/\Delta$ , where, again,  $\Delta$  is the minimum allowed increment for  $|x - x_0|$ . Additional analytic insight can be obtained by studying Problem (3.16).

The procedure for obtaining the Green's function outlined in Sect. 3.3 is referred to as "expansion of the Green's function in eigenfunctions" (or normal modes, see Chap. 6). A closed form for the Green's function with the given boundary conditions is available from the theory of differential equations, and can be used to find an alternative answer to (3.45) for  $y(x, t)$ . The interested reader is referred to Problem (3.17).



**Fig. 3.6.** Reflection of a pulse at a fixed end using the method of images

**Example 3.1.** What happens to the displacement of the string when a pulse such as the one approaching the left end of the string in Fig. 3.3b actually arrives at the fixed support?

*Solution.* Let us call the function which describes the original pulse on the string traveling to the left  $y(x, t) = Y(x + st)$ .

Suppose we consider the displacement  $y(x, t) = Y(x + st) - Y(st - x)$  with the second pulse in the neighborhood of  $x = -L/2$  at  $t = 0$ .  $y(x, t)$  is a perfectly acceptable solution. It

- a) satisfies the conditions at  $t = 0$  for  $L > x > 0$ ,
- b) satisfies the wave equation for the string,
- c) satisfies the boundary conditions at  $x = 0$  and  $x = L$  (the latter for a limited time only). At  $x = 0, y(0, t) = Y(st) - Y(st) = 0$ .

Since conditions (a–c) are the necessary and sufficient conditions for a solution, we have solved the problem. We note that  $Y(st - x)$  is a pulse traveling to the right, as shown. The effect of a fixed support is to produce a “reflected” wave.

In Problem (3.15) we use the method of images to validate the result obtained with a Green's function for a stretched string driven from the  $x = 0$  end. An experiment illustrating that the reflected wave leads to an overall standing wave is often demonstrated in elementary courses on wave physics. (However, swing to practical difficulties, it is sometimes difficult to make a convincing demonstration).

**Example 3.2.** Show in the usual way that the Green's function for a stretched string given in (3.41a,b,c) can also be written in closed form as:

$$G(x|x_0) = \frac{1}{k \sin(kL)} \begin{cases} \sin(kx) \sin[k(L - x_0)] & x < x_0 \\ \sin(kx_0) \sin[k(L - x)] & x > x_0 \end{cases}$$

*Solution.* The reader can easily show that  $G(x|x_0)$  satisfies the differential equation (3.33)

$$\frac{\partial^2 G(x|x_0)}{\partial x^2} + k^2 G(x|x_0) = -\delta(x - x_0) \quad L > x, x_0 > 0$$

when  $x \neq x_0$ , and that  $G$  vanishes at both  $x = 0$ , and at  $x = L$ . We treat what happens at  $x = x_0$  as an additional boundary condition, see Sect. 2.4. We must have:

$$\left. \frac{\partial G(x|x_0)}{\partial x} \right|_{x_0+\varepsilon} - \left. \frac{\partial G(x|x_0)}{\partial x} \right|_{x_0-\varepsilon} = -1$$

the integral over  $x$  of the second term on the left of (3.33) vanishing as usual. On the other hand, the derivation of  $G$  with respect to  $x$  has a discontinuity at  $x = x_0$ :

$$\begin{aligned} & \left. \frac{\partial G(x|x_0)}{\partial x} \right|_{x_0+\varepsilon} - \left. \frac{\partial G(x|x_0)}{\partial x} \right|_{x_0-\varepsilon} \\ &= 1/(\sin kL) [-\sin kx_0 \cos k(L-x_0) - \cos kx_0 \sin k(L-x_0)] \\ &= -1. \end{aligned}$$

In Problem 3.17 the reader is asked to verify that the closed form for  $G$  leads to the same result for the string driven at one end as is obtained using the method of Example 3.1, above. One may wonder how two such different forms for  $G$ , the closed form here, versus the formal expansion in normal modes (3.41) can be equal. The interested reader is, again, referred to Problem (3.17).

## Problems

**3.1** Find a general solution  $y(x, t)$  for a stretched string with initial displacement  $y_0(x) [=y(x, 0)]$ , and initial velocity  $v_0(x) [= \partial y(x, t)/\partial t |_{t=0}]$ . How does the development in Fig. 3.3 follow from your answer here?

*Hint:* Let  $y(x, t) = g_1(x + st) + g_2(x - st)$ , where  $g_1$  and  $g_2$  are to be found.

**3.2** Verify by direct integration (3.19a,b).

**3.3** Verify (3.21a) and (3.21b). Discuss the results contained in these equations.

*Hint:* see Example 1.1.

**3.4** A string of length  $L$ , with end points fixed, is vibrating purely in its second normal mode,  $n = 2$ . Its initial velocity is zero for all  $x$ .

a) Make a schematic drawing of the string at times

$$t = (1/8)(2\pi/\omega_2), \quad t = (1/4)(2\pi/\omega_2), \quad t = (5/8)(2\pi/\omega_2)$$

showing its displacement for all  $x$ .

b) Draw a diagram showing the displacement  $y$  as a function of time at the point  $x = (3/8)L$ , and at  $x = L/2$ .

c) Explain why what you have drawn corresponds to *standing waves*.

**3.5** A string of length of 1 meter is fixed at both its ends. How many normal modes does it have, if we count all wave vectors smaller than  $100 \text{ m}^{-1}$ , i.e.,  $k < 100 \text{ m}^{-1}$ ?

**3.6** The total kinetic energy of the vibrating string is given by the integral

$$\text{K.E.}(t) = 1/2 \int_0^L \varrho \left( \frac{\partial y}{\partial t} \right)^2 dx .$$

a) Why does this integral give the kinetic energy?

- b) The potential energy  $V(t)$  of a stretched string originates in the work done against the tension  $T$ , when as a result of its displacement, the string becomes longer. Show that

$$V(t) = \frac{T}{2} \int_0^L \left[ \frac{\partial y(x, t)}{\partial x} \right]^2 dx .$$

- c) Use our series expression (3.16b) to show that the total energy can be written as a sum over normal mode terms, with no cross-terms linking different modes.

- 3.7** a) The power  $P(x, t)$  transmitted at any point  $x$  by the string is given by the transverse force that the left side of the string exerts on the right side multiplied by the transverse velocity of the string at  $x$ . Show that for any wave  $y(x - st)$ ,  $P$  is given by

$$P(x, t) = sT(y')^2 ,$$

$$\text{where } y' = \frac{dy(\xi)}{d\xi} , \quad \xi \equiv x - st .$$

- b) Calculate the product of the energy density for the same wave as in part (a) (obtainable from the previous problem) and the wave velocity  $s$ . Compare the result with the expression for  $P(x, t)$  in part (a).

- 3.8** At time  $t = 0$  a string is deformed as shown in Fig. 3.7, and then released from rest.

- a) Express the displacement as an expansion in normal modes and calculate the coefficients  $A_n, B_n$ .

- b) The intensity of the sound emitted by the vibrating string at any of the normal frequencies is proportional to the square of the amplitude of the mode. Show that the intensity of sound corresponding to the fifth harmonic is less than 0.2 % of that corresponding to the fundamental mode.

*Hint:* Are any even harmonics present? And,

$$\int \frac{hx}{L/2} \sin ax dx = \frac{2h}{a^2 L} \sin ax - \frac{2hx}{La} \cos ax + C .$$

- c) Draw the string at  $t = 0$ , taking into account only harmonics up to the third. Discuss, also paying attention to the various signs ( $\pm$ ).

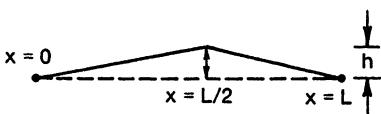


Fig. 3.7. Problem 3.8

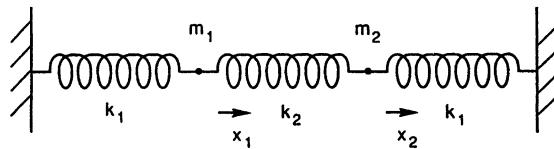


Fig. 3.8. Problem 3.9

**3.9** a)<sup>2</sup> Find the two normal modes for the coupled oscillator problem in Fig. 3.8.

*Hint:* In the equations of motion substitute

$$X_I(t) = X_I \cos \omega t \quad I = 1, 2$$

and find the two possible values of  $\omega$  which give nontrivial solutions for  $X_I$ .

Take  $m_1 = m_2 = m$ .

b) Show that when  $k_2 \ll k_1$ , the energy if initially carried by the left mass  $m_1$  will sledge back and forth between  $m_1$  and  $m_2$  with a period  $\pi/\beta$ ,  $\beta \equiv k_2/m$ ,  $m$  the common mass.

*Hint:* Express  $x_1(t)$  and  $x_2(t)$  in normal modes.

Take as initial condition:

$$\begin{aligned} x_1(0) &= A, & dx_1/dt|_{t=0} &= 0 \\ x_2(0) &= 0, & dx_2/dt|_{t=0} &= 0. \end{aligned}$$

Now calculate  $x_1(t)x_1^*(t)$ ,  $x_2(t)x_2^*(t)$ , i.e. the amplitudes squared.

c) Two identical pendula are coupled by having their suspension rods rigidly attached at the points of suspension to the two ends of a horizontal torsion bar. Give a qualitative analysis of what happens if the left bob is pulled to one side, (i.e. perpendicular to the bar) and released from rest, while simultaneously the right bob is released from its equilibrium position at the bottom.

This experiment makes for a good laboratory demonstration.

**3.10** A wire of length 1-1/2 meters weighs 2.5 grams. What mass in grams must be suspended from its end so that resonance with the third normal mode is achieved when a driving electromagnet is operated at 60 Hz?

**3.11** The general way of expressing an arbitrary function  $y(x)$  periodic in the interval 0 to  $L$  is by means of the complete, periodic set of functions

$$y(x) = \sum_{n=1}^{\infty} A_n \sin \frac{2n\pi x}{L} + \sum_{n=0}^{\infty} B_n \cos \frac{2n\pi x}{L}.$$

How is it, then, that the set  $\sin(n\pi x/L)$ ,  $n = 1, 2, 3 \dots$  has been cited as complete, admittedly for fixed end points?

*Hint:* Think about the larger interval 0 to  $2L$ , and consider a function of suitable symmetry.

---

<sup>2</sup> Problem 3.9 is important for Chap. 6 on wave mechanics.

- 3.12** a) Show that the series for  $y(x)$  in Problem 3.11 can be written as

$$y(x) = \sqrt{2\pi}/L \sum_{n=-\infty}^{\infty} a(k_n) e^{ik_n x}$$

where  $k_n = 2\pi n/L$ . Find  $a_n$  in terms of  $A_n$  and  $B_n$ .

- b) Show that

$$a_n(k_n) = 1/\sqrt{2\pi} \int_{-L/2}^{+L/2} y(x) e^{-ik_n x} dx$$

where  $x$  extends from  $-L/2$  to  $L/2$ , instead of from 0 to  $L$ .

- c) What integral does  $y(x)$  in (a) approach as  $L \rightarrow \infty$ ? What is the corresponding equation for  $a(k)$ ?

- 3.13** In a sound wave in a gas traveling in the  $x$  direction the excess pressure  $p(x, t)$  is proportional to the compression  $\partial y(x, t)/\partial x$ , viz.  $p(x, t) = -B\partial y(x, t)/\partial x$ . Here  $y$  is the average molecular displacement from equilibrium, and  $B$  is the “adiabatic bulk modulus of elasticity”. If sound waves of fixed frequency generated by a tuning fork are sent down a glass column whose length can be adjusted, then resonances set in.

- a) By considering the forces on and acceleration of a cross-sectional slice of gas of thickness  $\Delta x$ , show that

$$s^2 \frac{\partial^2 y}{\partial x^2} = \frac{\partial^2 y}{\partial t^2}$$

where  $s^2 = B/\varrho$ ,  $\varrho$  is the density.

- b) Find the set of normal modes  $y_n(x, t)$  for the average displacement of air molecules from equilibrium taking into account the boundary conditions of no excess pressure on top of the column, and no average displacement at the closed-off bottom. Are the modes *longitudinal* (displacement in the direction of propagation), or are they transverse, as the waves on a string?

- c) Discuss the completeness of the set of modes you have suggested following the method of Problem 3.9.

*Hint:* Consider the range  $-2L < x < 2L$ .

- d) Without actually doing the calculation, explain how you would go about deriving a result analogous to (3.28), and suggest what the result might look like, assuming that the system is being driven from somewhere near the head of the air column. At what frequencies do we find resonances?

- 3.14** A force  $F(t) = F_0 \cos \omega_0 t$  is applied vertically to the  $x = 0$  end of a long string of density  $\varrho$  and horizontal tension  $T$ . (A long string corresponds to the situation where any reflected signal is either too small or comes too late for consideration.)

- a) Why do we call the expression  $y = -y_0 \sin(\omega_0 x/s - \omega_0 t)$  a *traveling wave*? What is the physical interpretation of the relation  $\lambda f = s$  for this wave?

*Hint:* How many waves pass a fixed observer in unit time?

- b) Under the assumption that a traveling wave originates at  $x = 0$  which is the driven end of the string, show that the ratio  $F(t)/[\partial y(0, t)/\partial t]$  is given at all times by

$$\frac{F(t)}{\partial y/\partial t} = \frac{T}{s}$$

where  $s$  is the wave speed. This ratio is the mechanical impedance (electrical analogue  $\varepsilon/I$ ) characteristic of the string. What happens qualitatively to the travelling wave if some length of the string is tied to a stretched string with the same characteristic impedance, with a different characteristic impedance?

*Hint:* See Problem 6.11.

- c) Use the principle of superposition to explain why for a long force-free string extending from  $x = 0$  we can write

$$\begin{aligned} & \frac{\partial y(x, t)}{\partial t} \\ &= \frac{1}{\sqrt{2\pi}} \int_0^\infty [A(\omega) \cos(\omega t - \omega x/s) + B(\omega) \sin(\omega t - \omega x/s)] d\omega \quad x > 0. \end{aligned}$$

- d) If at the  $x = 0$  end

$$\frac{\partial y(0, t)}{\partial t} = 1/2 \frac{F_0 s}{T} (e^{i\omega_0 t} + e^{-i\omega_0 t}) e^{-t^2/4n},$$

show that

$$A(\omega) = \sqrt{2\pi} \left( \frac{F_0 s}{T} \right) [g_n(\omega_0 - \omega) + g_n(\omega_0 + \omega)]$$

with

$$g_n(x) \equiv \sqrt{n/\pi} e^{-nx^2},$$

$$B(\omega) = 0.$$

- e) Treating  $\exp[i(\omega t - \omega x/s)]$  as a good function, use parts (c) and (d) to find  $\partial y(x, t)/\partial t$ , and  $y(x, t)$  in the limit  $n \rightarrow \infty$ .

- f) Use parts (c) and (d) to find  $\partial y(x, t)/\partial t$  without letting  $n \rightarrow \infty$ , and discuss your result.

- 3.15** a) The  $x = 0$  end of a string under tension and fixed at  $x = L$  is driven so that  $y(0, t) = A \cos \omega t$ . Use the method of images to show that:

$$y(x, t) = \frac{A}{\sin(\omega L/s)} \sin [\omega(L - x)/s] \cos \omega t$$

*Hint:* Start by letting

$$y(x, t) = \alpha \sin [k(L - x) + \omega t + \delta] + \beta \sin [k(x - L) + \omega t + \delta']$$

an incoming and reflected wave, with no loss of generality.  $\delta, \delta'$  are initially unknown phases. Compare your result with the statement that, if one end of the string is harmonically driven while the other end is fixed, a standing wave forms on the string. What happens as the driving frequency approaches a natural frequency, i.e. the frequency of a normal mode?

- b) Show that the complex impedance of the string,  $Z$  [the ratio of the applied force to the transverse velocity as in Problem (3.14b)] is given by:

$$Y = Z^{-1} = -i/(s\varrho) \tan(\omega L/s)$$

$Y$ , the reciprocal of  $Z$  is called the admittance. Does the string absorb power? Discuss your result for  $Z$  briefly, i.e. as a function of the driving frequency  $\omega$ . When  $Z$  is infinite, what is the  $x = 0$  end of the string doing during vibration? Could one build an oscillator from a vibrating string?

N.B. The electric analogue of the vibrating string is a capacitor and inductor in parallel. See Problem 2.11.

- 3.16** a) The  $x = 0$  end of a string under tension and fixed at  $x = L$  is driven so that  $y(0, t) = A \cos \omega_0 t$ . Find  $y(x, t)$  by letting  $z(x, t) = y(x, t) + [A(x - L)/L] \cos \omega_0 t$ , and using the method of Sect. 3.4.

- b) Check that your result agrees with (3.45).

*Hint:*  $(1 - x/L)$  can be represented as:

$$1 - \frac{x}{L} = \sum_{n=1}^{\infty} \frac{2}{n\pi} \sin\left(\frac{n\pi x}{L}\right) \quad L > x > 0.$$

The right side of this equation is not defined at  $x = 0$ . It is this contribution to  $y(x, t)$  that leads to the convergence difficulties with (3.45). See also the discussion at the end of Sect. 3.5.

- c) Compare and contrast your result with the statement that if one end of a string is driven at a natural frequency while the other end is fixed, the corresponding standing wave forms on the string.

- 3.17** a) The closed form for the Green's function of the stretched string given in Example 3.2 is:

$$G_k(x|x_0) = \frac{1}{k \sin(kL)} \begin{cases} \sin(kx) \sin[k(L - x_0)] & x < x_0 \\ \sin(kx_0) \sin[k(L - x)] & x > x_0 \end{cases}$$

Use this result, along with the magic rule (3.40), to verify the expression for  $y(x, t)$  found in Problem 3.15 for the string driven from the  $x = 0$  end.

- b) Show that the Green's function given in (3.41), i.e.

$$G_k(x|x_0) = \sum_{n=1}^{\infty} \frac{u_n(x)u_n(x_0)}{k_n^2 - k^2} \quad k \neq k_n, x \neq x_0, L > x, x_0 > 0$$

is equal to the closed form result, above, in Part (a) for all physical  $x$  and  $x_0$ . Liouville's theorem in complex variable analysis states that a function of  $z$  which is analytic for all finite values of  $z$  and is bounded everywhere is a constant. It follows that two functions with the same singularities in the complex plane can differ only by a constant. Here the reader must show that if one of the forms of  $G_k(x|x_0)$  goes to infinity in the complex  $k$  plane as  $A/(k-\alpha)$  as  $k \rightarrow \alpha$ , the other form for  $G_k(x|x_0)$  goes to infinity as  $B/(k-\beta)$ , with  $A = B$ ,  $\alpha = \beta$ .

Why must we check that both  $G$ 's go to zero as  $k$  goes to  $\pm\infty$ , and, also  $k$  goes to  $\pm i\infty$ ?

**3.18** Give a plausibility argument for the first term  
 $-G(x|x_0)\partial u(x_0)/\partial x_0|_{x_0=0}$  in the magic rule (3.40).

*Hint:* In general with a source  $S\delta(x-\Delta)$  at a point  $\Delta$ :

$$u(x) = G(x|\Delta)S$$

If  $du(x)/dx|_{x=0} \equiv \alpha$  is specified, then we can obtain a solution with (3.40) by choosing  $\partial G(\Delta|x_0)/\partial x_0 = 0$  at  $x_0 < \Delta$ . Reciprocity yields:  $\partial G(x|\Delta)/\partial x = 0$  at  $x < \Delta$ .

We can, now, find the source strength  $S$  at  $x_0 = \Delta$  of the simulating force from the equation:

$$S\partial G(x|\Delta)/\partial x|_{x=\Delta+\varepsilon} = \alpha$$

See Example 3.2 for inspiration in finding the unknown derivative in this last equation, thereby obtaining  $S$ . (This  $S$  will produce the given initial condition for  $x > \Delta$ , rather than having the derivative vanish, and at the same time will yield the very motion as the first term in the magic rule.)

**3.19** Taking the full complex result obtained in Sect. 3.4 for  $C_n$  show that for a long string, and when  $b \ll \omega$ , the driven solution (3.27) reduces to

$$y(x,t) = \frac{F_0}{\varrho s} \operatorname{Re} \left\{ \frac{-i}{\omega - ib/2} \exp \left[ i\omega \left( t - \frac{x-x_0}{s} \right) \right] \exp \left[ \frac{-(x-x_0)b}{2s} \right] \right\}$$

when  $x > x_0$

$$= \frac{F_0}{\varrho s} \operatorname{Re} \left\{ \frac{-i}{\omega - ib/2} \exp \left[ i\omega \left( t + \frac{x-x_0}{s} \right) \right] \exp \left[ \frac{(x-x_0)b}{2s} \right] \right\}$$

when  $x < x_0$ .

These are two damped traveling waves originating at  $x = x_0$  heading in opposite directions. Here a long string means that  $L \gg s/b$  and that  $L\omega/s \gg 1$ , where  $\omega$  is the driving frequency, and that the driving point  $x_0$  is likewise far from either end.

Note that for small  $b/\omega$

$$\int_{-\infty}^{\infty} \frac{e^{ivP} dv}{v^2 + ib\omega - \omega^2} \simeq \begin{cases} \frac{\pi i}{(\omega - ib/2)} \exp[-P(i\omega + b/2)] & \text{when } P > 0, \\ \frac{\pi i}{(\omega - ib/2)} \exp[P(i\omega + b/2)] & \text{when } P < 0. \end{cases}$$

## Further Reading

W.C. Elmore, M.A. Heald: *Physics of Waves* (Dover, New York 1985)

P.R. Wallace: *Mathematical Analysis of Physical Problems* (Dover, New York 1984)

P.M. Morse, K.V. Ingard: *Theoretical Acoustics* (McGraw-Hill, New York 1968)

P.M. Morse, H. Feshbach: *Methods of Theoretical Physics* (McGraw-Hill, New York 1953)

## 4. Electromagnetic Waves

**Abstract.** The development of one dimensional waves in Chap. 3 is applied to the six dimensional system of electromagnetic fields. The integral form of Maxwell's equations, assumed to be familiar to the reader from introductory treatments, begins this chapter. Mathematical definitions and the physical significance of the vector operators of gradient, divergence, and curl are given. In the appendix to this chapter the reader will find proofs establishing the equivalence for these operators of the definitions and physical interpretations. The differential form of Maxwell's equations follows. Beginning with free space, we then show the Maxwell equations leading to wave equations for the electric and magnetic fields. Solutions of these wave equations give us the propagation of electromagnetic waves, i.e., light in the visible region, etc. The solutions can be modified to apply to closed systems, cavities, in analogy to a finite string with fixed ends. Again, carrying over from Chaps. 2 and 3 we have normal modes, natural motion, and, Problem 4.21, driven motion and resonance. Finally, we solve Maxwell's equations for the general case, i.e., with distributed sources, charge and current densities, pervading the space. The Helmholtz theorem, which specifies what it is that determines vector fields, is given, and used to find the electric and magnetic fields in terms of the sources. We, then, apply the general solutions to calculating the power radiated by a dipole antenna. In the last section we show how the Green's function technique provides a systematic solution to the driven wave equation.

### 4.1 Maxwell's Equations in Integral Form

The evolution of magnetic and electric fields in space and time has much in common with the mechanical motion of distributed systems of the previous chapter. The fields will also be found to satisfy certain wave equations. The *sources* of the mechanical disturbances are applied forces, and they commonly appear on the right side of the wave equations; we have called them the driving terms. The sources of the electromagnetic fields are the electric charges  $q$  and the electric current densities  $\mathbf{j}$ , and they, also, are found on the right side of the corresponding wave equations. Electromagnetism is complicated

by the fact that there are two fields, the electric field  $\mathbf{E}$ , and the magnetic field  $\mathbf{B}$ , and that these are coupled together.

The reader is assumed to have some familiarity with the integral form of the laws of electromagnetism:

$$\oint \mathbf{E} \cdot d\mathbf{S} = 1/\epsilon_0 \oint \varrho dv \quad (4.1)$$

$$\oint \mathbf{E} \cdot d\mathbf{l} = - \int \partial \mathbf{B} / \partial t \cdot d\mathbf{S} \quad (4.2)$$

$$\oint \mathbf{B} \cdot d\mathbf{S} = 0 \quad (4.3)$$

$$\oint \mathbf{B} \cdot d\mathbf{l} = \mu_0 \int (\mathbf{j} + \epsilon_0 \partial \mathbf{E} / \partial t) \cdot d\mathbf{S}. \quad (4.4)$$

These four relations are the integral form of *Maxwell's equations*. They enable us to find the vector fields  $\mathbf{E}(\mathbf{r}, t)$  and  $\mathbf{B}(\mathbf{r}, t)$  for any given charge per unit volume,  $\varrho(\mathbf{r}, t)$ , and any given distribution of electric current per unit area,  $\mathbf{j}(\mathbf{r}, t)$ . We shall show how to find the fields, i.e., to solve the equations, in the course of this chapter.

The first equation (4.1) is Gauss's Law. It states that the integral of the electric field over any arbitrary closed area, the so-called electric flux, is equal to the total charge enclosed within the volume bounded by that same area (divided by  $\epsilon_0$ ). The second equation, Faraday's Law, tells us that the line integral of the electric field around any closed path is equal to minus the time rate of change of the magnetic flux, i.e., the magnetic field integrated over the area enclosed by that path. The latter two equations contain integrals of the same nature as the two preceding ones. In (4.3) we see that there are no magnetic sources analogous to the electric charge density. Equation (4.4) is Ampere's Law as modified by Maxwell.  $\epsilon_0$  and  $\mu_0$  are called the permittivity and permeability of free space;  $\epsilon_0 = 8.85 \times 10^{-12}$ , and  $\mu_0 = 4\pi \times 10^{-7}$ , in the mks system of units.

As the reader knows, these equations can be used directly only in situations with special symmetry. For example, with (4.1) we can always find the electric field of a charged plate of infinite extent. Similarly, (4.4) enables one to find at once the magnetic field of an infinitely long current carrying conductor.

## 4.2 Maxwell's Equations in Differential Form

Maxwell's equations in differential form are in terms of the operators *gradient*, *divergence*, *curl*. These operate on vector fields in a prescribed manner, and the result of each operation has a physical interpretation. We will give both the manner of operation and the interpretations here. These interpretations are not obvious. In the appendix we outline proofs in the simpler situation

of two dimensions. The interested reader may see what sort of considerations are involved.

In cartesian coordinates we define the vector operator “ $\nabla$ ” as

$$\nabla \equiv (\partial/\partial x)\mathbf{i} + (\partial/\partial y)\mathbf{j} + (\partial/\partial z)\mathbf{k} \quad (4.5)$$

with  $\mathbf{i}, \mathbf{j}, \mathbf{k}$  the unit vectors in the  $x, y, z$  direction.

*Gradient:* The definition is

$$\text{grad}\psi(x, y, z) \equiv \nabla\psi = (\partial\psi/\partial x)\mathbf{i} + (\partial\psi/\partial y)\mathbf{j} + (\partial\psi/\partial z)\mathbf{k}, \quad (4.6a)$$

where  $\psi$  is a scalar function of  $x, y, z$  on which we are to operate. The gradient creates out of a scalar field  $\psi(x, y, z)$  the vector field  $\nabla\psi$ . If  $\mathbf{c}$  is a unit vector in some given direction, then the physical interpretation of  $\nabla\psi$  is that  $\mathbf{c} \cdot \nabla\psi$  gives the rate of change of  $\psi$  in the direction of  $\mathbf{c}$ . Noticing that

$$(\mathbf{c} \cdot \nabla\psi) = |\nabla\psi| \cos \theta$$

where  $\theta$ , the angle between the two vectors, goes to zero when the unit vector  $\mathbf{c}$  is aligned with  $\nabla\psi$ , we see that  $\nabla\psi$  points in the direction of the greatest rate of change of  $\psi$ , and its magnitude is that greatest rate of change. In electrostatics the relation between the electric field and the scalar potential  $\phi$  is  $\mathbf{E} = -\nabla\phi$ , as the reader may know.

*Divergence:*

$$\text{div}\mathbf{F}(x, y, z) \equiv \nabla \cdot \mathbf{F} = \partial F_x/\partial x + \partial F_y/\partial y + \partial F_z/\partial z. \quad (4.6b)$$

We formally take the vector dot product between  $\nabla$  and  $\mathbf{F}$ , and create out of the vector field  $\mathbf{F}(x, y, z)$  a scalar field  $\nabla \cdot \mathbf{F}$ . The divergence measures the net outward flux of the vector field  $\mathbf{F}$  from a vanishingly small volume, per unit volume. That is

$$\nabla \cdot \mathbf{F} = \frac{\oint \mathbf{F} \cdot d\mathbf{S}}{\oint dv} \quad \text{as} \quad \oint dv \rightarrow 0. \quad (4.6c)$$

The term “flux” for the surface integral of the vector field has its origin in fluid mechanics where the field is the velocity and the flux is then the total fluid flow through the surface per unit time.

*Curl:*

$$\begin{aligned} \text{curl}\mathbf{F}(x, y, z) \equiv \nabla \times \mathbf{F} &= [(\partial/\partial x)\mathbf{i} + (\partial/\partial y)\mathbf{j} + (\partial/\partial z)\mathbf{k}] \\ &\times [F_x\mathbf{i} + F_y\mathbf{j} + F_z\mathbf{k}]. \end{aligned} \quad (4.6d)$$

The curl is obtained by formally calculating the vector crossproduct of del and the vector field  $\mathbf{F}(x, y, z)$  on which it operates. A vector field again yields a vector field. The physical interpretation of the curl is that if we take the line integral of a vector field  $\mathbf{F}$  around the contour enclosing a vanishingly small area  $\Delta\mathbf{S}$ ,  $\oint \mathbf{F} \cdot d\mathbf{l}$ , then in the limit as the magnitude of the area  $|\Delta\mathbf{S}|$  goes to zero, the curl satisfies

$$(\nabla \times \mathbf{F}) \cdot \Delta \mathbf{S} = \oint \mathbf{F} \cdot d\mathbf{l} . \quad (4.6e)$$

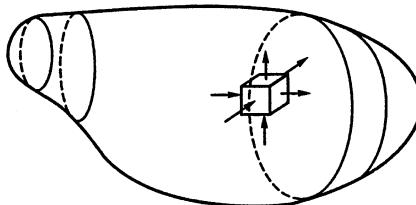
Here  $\Delta \mathbf{S}$  points perpendicularly to the plane of integration in the usual right-hand rule manner. It follows, recollecting our discussion of the gradient, that the direction of the curl is perpendicular to that plane of integration oriented to yield the greatest value for the line integral per unit area. The magnitude of the curl is the line integral around  $\Delta \mathbf{S}$  when so oriented, divided by  $|\Delta \mathbf{S}|$  in the limit that  $|\Delta \mathbf{S}|$  goes to zero.

The curl can be said to give the “circulation” per unit area at a given point in the vector field. It measures the rate of twisting of lines of force, or of flow, the “vorticity” of the field on which it is acting. If the field is the velocity of a fluid, a nonvanishing curl implies rotation of the fluid.

To someone exposed to them for the first time, these three definitions may seem like much to assimilate. We believe that if the readers follow the rest of the development of this chapter and study the examples and problems, they will gradually acquire a sufficient mastery of these vector operators.

There exist two basic theorems – one concerning the divergence, discovered by Gauss, the other, concerning the curl, discovered by the English physicist G. Stokes (1819–1903). Consider the integral  $\oint \nabla \cdot \mathbf{F} dv$  taken over some arbitrary closed volume such as the one shown in Fig. 4.1,  $\mathbf{F}(x, y, z)$  is a continuous vector field. From the definition of the divergence we know that an element of the integrand,  $\text{div } \mathbf{F} dv$  ( $dv$  being the element of volume), equals the net flux issuing from the element, i.e., flux out less flux in. The flux out from any element either enters the next element or, finally, if the element is at the surface, it issues. Accordingly, we have Gauss’s theorem:

$$\int \nabla \cdot \mathbf{F} dv = \int F \cdot d\mathbf{S} . \quad (4.7)$$



**Fig. 4.1.** The net flux issuing from the small volume  $\Delta v$  is  $[\nabla \cdot \mathbf{F}(x, y, z)]\Delta v$ . If we integrate over the entire closed arbitrary volume, we shall be left with the net flux issuing from its outside surface – Gauss’s theorem

In Fig. 4.2 we show an arbitrary surface bound by some particular contour. We want to consider the following integral taken over the surface bounded by the contour:

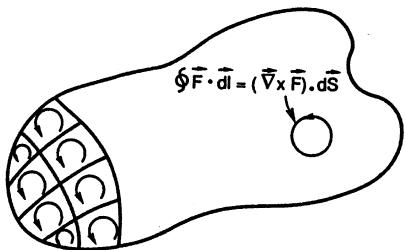
$$\int (\nabla \times \mathbf{F}) \cdot d\mathbf{S} .$$

We already know that the element  $(\text{curl } \mathbf{F} \cdot d\mathbf{S})$  represents the line integral of the vector field  $\mathbf{F}$  around the little contour binding  $d\mathbf{S}$ . If we begin to add

the contributions to the integral from adjoining elements of area then the line integral over every internal contour will cancel, as there will always be line integrands in opposing directions. All that will be left is the line integral of  $\mathbf{F}$  around the contour binding the entire area. Accordingly, we have Stokes's theorem, which reads

$$\int (\nabla \times \mathbf{F}) \cdot d\mathbf{S} = \oint \mathbf{F} \cdot dl. \quad (4.8)$$

The integral of the curl of a vector field over an arbitrary surface is equal to the line integral of the vector field around the contour binding the surface.



**Fig. 4.2.** The line integral around one of the little contours is given by  $[\nabla \times \mathbf{F}(x, y, z)] \cdot d\mathbf{S}$ . If we integrate over the entire surface, then line integrals over internal contours will cancel, giving us the line integral of  $\mathbf{F}$  over the outside contour – Stokes' theorem

We now use these theorems to recast the Maxwell equations (4.1–4) into their more accessible form as differential equations. We can write the first Maxwell equation (4.1) as

$$\nabla \cdot \mathbf{E} = \rho/\epsilon_0. \quad (4.9)$$

Obviously we can go back from here to (4.1) by Gauss's theorem simply by integrating both sides of (4.9) over an arbitrary volume. It is the *arbitrary* ranges of all integrations in the integral form of the equations that gives them their power. By choosing two volumes in (4.1) only slightly removed from one another and subtracting, we can extract the Maxwell equation in the differential form (4.9) directly from (4.1). The two forms of the equation are exactly equivalent.

Similarly, the use of Stokes's theorem, (4.8), leads to a differential form for (4.2):

$$\nabla \times \mathbf{E} = -\partial \mathbf{B} / \partial t. \quad (4.10)$$

And having gone this far, we easily finish by rewriting (4.3) and (4.4):

$$\nabla \cdot \mathbf{B} = 0, \quad (4.11)$$

$$\nabla \times \mathbf{B} = \mu_0 [\mathbf{j} + \epsilon_0 \partial \mathbf{E} / \partial t]. \quad (4.12)$$

Equations (4.9–4.12) are the differential form of Maxwell's equations. These are the equations most generally used in more advanced treatments of electromagnetic theory.

## 4.3 Plane Electromagnetic Waves in Free Space

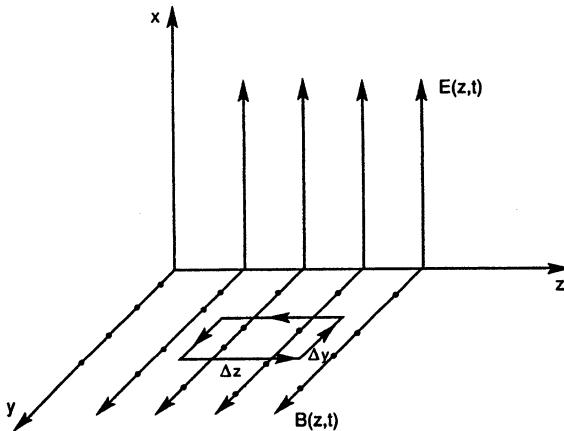
In Chap. 3, Waves on Stretched Strings, we learned to solve two kinds of problems, that of the subsequent development of a system which had been subjected to a certain kind of initial disturbance, the so-called initial value problem, and second, the problem of determining the response of the system to a prescribed set of continuously applied forces, the forced motion of the system. We shall now study the analogous problems for electric and magnetic fields.

We have already seen that the equivalents to applied forces in mechanical systems are the sources of the fields, electric charges and electric currents. In this section, as a beginning, we will analyze the situation when there are no sources, only a disturbance, originally produced either at an earlier time, or at a different place. Accordingly, we have free space. (The presence of matter generally implies the existence of some charges and currents, see next chapter.) For simplicity we shall restrict ourselves to cartesian coordinates. Although we wish to work in three dimensions, it is usual to analyze problems in any coordinate system by first treating disturbances which vary in only one of the coordinates; such disturbances do not have any spatial variations whatever over surfaces that, in the cartesian system, are planes. For example, for a disturbance in the  $z$ -direction one has the planes with  $z$  equal to various constants, which are the  $x$ - $y$  plane or the various planes parallel to it. In this section we shall consider, then, plane electromagnetic waves in free space.

We are seeking plane solutions for  $\mathbf{E}$  and  $\mathbf{B}$  to the Maxwell equations, solutions  $\mathbf{E}(z, t)$ ,  $\mathbf{B}(z, t)$  that depend on  $z$  and the time  $t$ , but not on  $x$  or  $y$ . It is not obvious that we can find such solutions, but we shall show that we can, and point out how they may be extended to form a *complete set*. Thus, we shall have the desired basis for doing any electromagnetic problem in free space despite our having started with restrictions.

The plane solutions we are seeking will be *transverse*. A transverse field may be taken as one which has zero divergence. It follows from (4.9),  $\text{div } \mathbf{E} = \varrho/\epsilon_0$ , that in free space, where the charge density  $\varrho$  is zero,  $\nabla \cdot \mathbf{E} = 0$ , and we always have  $\nabla \cdot \mathbf{B} = 0$ . Alternatively, we mean by a transverse solution that  $\mathbf{E}$  and  $\mathbf{B}$  will lie completely in the planes perpendicular to the direction of propagation, the  $z$ -direction. In the Problems we shall show that the two definitions of transverse are equivalent.

We still have no knowledge as to the direction of the  $\mathbf{E}$  and  $\mathbf{B}$  vectors in the  $x$ - $y$  planes. It turns out, as we shall see, that there will be no loss of generality if we initially assume that the  $\mathbf{E}$  field, and, in consequence, also the time derivative of  $\mathbf{E}$ ,  $\partial\mathbf{E}/\partial t$ , points in the same direction for all  $z$  and  $t$ . We have the situation shown in Fig. 4.3, where we have taken the  $x$ -direction for  $\mathbf{E}$  and  $\partial\mathbf{E}/\partial t$ . In this figure we show the projections of the electric lines of force on the  $x$ - $z$  plane as vertical lines, and on the  $y$ - $z$  plane by dots signifying the field emerging from the latter plane in a perpendicular direction.



**Fig. 4.3.** Transverse EM wave. The dots signify the emerging electric field out of the  $y$ - $z$  plane. We calculate  $\oint \mathbf{B} \cdot d\mathbf{l}$  around the rectangle shown, and equate it to  $\mu_0$  times the displacement current  $\int \epsilon_0 \cdot \partial \mathbf{E} / \partial t \cdot d\mathbf{S}$  threading this circuit

Let us apply Maxwell's equation in integral form (4.4) (extension of Ampere's Law) to this situation. We choose a small rectangle in the  $y$ - $z$  plane, as shown in Fig. 4.3, and integrate the magnetic field around its boundary. Since  $\mathbf{B}$  depends only on  $z$  and not on  $y$  it is obvious that the contributions to the line integral from the two segments parallel to the  $z$ -axis cancel each other. For the remaining two segments going counter clock-wise, as directed by the right-hand rule for a current in the positive  $x$ -direction, we have

$$\oint \mathbf{B} \cdot d\mathbf{l} = -B_y(z + \Delta z, t)\Delta y + B_y(z, t)\Delta y .$$

In the absence of current density, (4.4) is

$$\oint \mathbf{B} \cdot d\mathbf{l} = \mu_0 \epsilon_0 \int \partial \mathbf{E} / \partial t \cdot d\mathbf{S} .$$

For our infinitesimal rectangle, (4.4) combined with our result for the line integral of  $\mathbf{B}$  gives

$$-[B_y(z + \Delta z, t) - B_y(z, t)]\Delta y = \mu_0 \epsilon_0 \frac{\partial E_x(z, t)}{\partial t} \Delta z \Delta y .$$

Dividing both sides by  $\Delta z$  and going to the limit  $\Delta z \rightarrow 0$  we find

$$-\frac{\partial B_y(z, t)}{\partial z} = \mu_0 \epsilon_0 \frac{\partial E_x(z, t)}{\partial t} . \quad (4.13)$$

We could get more information by integrating around rectangles in the other two planes. Actually we can get the relations we are seeking directly from the differential form of the Maxwell equation, that is, from (4.12). In cartesian components we have

$$\begin{aligned}\nabla \times \mathbf{B} &= \left[ \left( \frac{\partial}{\partial x} \right) \mathbf{i} + \left( \frac{\partial}{\partial y} \right) \mathbf{j} + \left( \frac{\partial}{\partial z} \right) \mathbf{k} \right] \times [B_x(z, t)\mathbf{i} + B_y(z, t)\mathbf{j}] \\ &= \left( \frac{\partial}{\partial z} \right) \mathbf{k} \times [B_x(z, t)\mathbf{i} + B_y(z, t)\mathbf{j}] \\ &= \left( \frac{\partial B_x(z, t)}{\partial z} \right) \mathbf{j} - \left( \frac{\partial B_y(z, t)}{\partial z} \right) \mathbf{i}.\end{aligned}$$

Remembering that  $\partial \mathbf{E}/\partial t$  is in the  $x$ -direction, and that  $\mathbf{j} = 0$ , (4.12) reads

$$\nabla \times \mathbf{B} = \mu_0 \epsilon_0 \left( \frac{\partial E_x(z, t)}{\partial t} \right) \mathbf{i}.$$

Or, combining with our result for  $\nabla \times \mathbf{B}$ , we have

$$\left( \frac{\partial B_x(z, t)}{\partial z} \right) \mathbf{j} - \left( \frac{\partial B_y(z, t)}{\partial z} \right) \mathbf{i} = \mu_0 \epsilon_0 \left( \frac{\partial E_x(z, t)}{\partial t} \right) \mathbf{i}.$$

The  $x$ -component of this last equation just gives us again (4.13), as it should. The  $y$ -component of the equation tells us

$$B_x(z, t) = 0. \quad (4.14a)$$

If boundary conditions so warrant, we can always add in fields that are spatially constant.

Let us now process the differential Maxwell equation (4.10) in the same way as we have (4.12). By analogy, remembering that  $\mathbf{E}$  has only an  $x$ -component, we find

$$\left( \frac{\partial E_x(z, t)}{\partial z} \right) \mathbf{j} = - \left( \frac{\partial B_x}{\partial t} \right) \mathbf{i} - \left( \frac{\partial B_y}{\partial t} \right) \mathbf{j} - \left( \frac{\partial B_z}{\partial t} \right) \mathbf{k}.$$

Actually, we already know that  $B_x$  does not exist and that  $B_z$  vanishes, since both  $\mathbf{E}$  and  $\mathbf{B}$  are transverse. Here we see that  $B_z$  would have to be independent of time. Since we are not interested here in any stationary situation that might be superimposed on our time-dependent solutions, we also have that

$$B_z(z, t) = 0. \quad (4.14b)$$

By looking at both (4.14a) and (4.14b) we learn that the magnetic field, like the electric field, is confined to one direction, that of the  $y$ -axis. By looking at the  $y$ -component of our last relation, we get a second equation to go with (4.13),

$$\mu_0 \epsilon_0 \frac{\partial E}{\partial t} = - \frac{\partial B}{\partial z}, \quad (4.13)$$

$$\frac{\partial E}{\partial z} = - \frac{\partial B}{\partial t}, \quad (4.15)$$

where we have dropped the  $x$  subscript on  $E$ , and the  $y$  subscript on  $B$ .

It is easy to uncouple these equations. By differentiating both sides of the first equation with respect to  $t$ , and of the second with respect to  $z$ , and then subtracting, we find

$$\mu_0 \epsilon_0 \frac{\partial^2 E}{\partial t^2} - \frac{\partial^2 E}{\partial z^2} = 0. \quad (4.16a)$$

We can also differentiate the first with respect to  $z$ , and the second with respect to  $t$ , and add. Then we find

$$\mu_0 \epsilon_0 \frac{\partial^2 B}{\partial t^2} - \frac{\partial^2 B}{\partial z^2} = 0. \quad (4.16b)$$

Now each of these equations is precisely of the same form as the wave equation we derived in the previous chapter for the displacement of the stretched string. We, therefore, have solutions, for example, of the form

$$E = E_0 \sin(kz \pm \omega t), \quad (4.17a)$$

$$B = B_0 \sin(kz \pm \omega t). \quad (4.17b)$$

$E_0$  and  $B_0$  are amplitudes for the disturbance, and  $k$  and  $\omega$  are the wave vector and angular frequency, respectively. We note that we have taken the same wave vector and frequency for  $\mathbf{E}$  and  $\mathbf{B}$ , and are taking them in phase. These restrictions are necessary to satisfy (4.13) and (4.15). We must also have

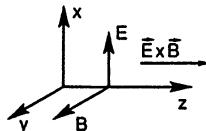
$$E_0 = cB_0. \quad (4.17c)$$

From the wave equations we know the speed of the wave; it is given by

$$c = 1/\sqrt{\mu_0 \epsilon_0}. \quad (4.18)$$

The reader may well be familiar with the fact that when the speed of light was measured, it was found to agree very well with the value one calculates from (4.18) using measured values of the permittivity and permeability of free space. This agreement at once elucidated the nature of light. The entire development, the analytical part credited above all to James Clerk Maxwell (1831–1879), and the most successful measurements (in 1880) to Albert A. Michelson (1852–1931), can be considered as possibly the greatest discovery in physics to date, certainly of the 19th century. All of electromagnetic radiation, from gamma rays at the high frequency end of the spectrum to long radio waves at the low end, satisfy the Maxwell equations, and propagate with the same velocity  $c$  in free space. The understanding of all these phenomena originates from the more limited laws first discovered by Coulomb, Ampere, and Faraday.

The picture we have so far of light is that of time- and space-varying electric and magnetic fields as in (4.17), with the electric vector being fixed in one direction (our  $x$ -axis), the magnetic vector fixed at right angles to it (our  $y$ -axis), and the light propagating in the direction mutually perpendicular to these two (Fig. 4.4).



**Fig. 4.4.** The direction of propagation, given by  $\mathbf{E} \times \mathbf{B}$ , the electric field  $\mathbf{E}$ , and the magnetic field  $\mathbf{B}$  form a mutually orthogonal set. The Poynting vector  $\mathbf{S} = (\mathbf{E} \times \mathbf{B}/\mu_0)$  gives the intensity (see Problems)

The reader may know that even in free space electric and magnetic fields carry energy, with energy densities  $(1/2)\epsilon_0 E^2$  and  $B^2/(2\mu_0)$ , respectively. As a result, one can show (see Problems) that the intensity (energy/unit time – unit area) of the electromagnetic wave is given by the so-called Poynting vector  $\mathbf{S}$  (after John Henry Poynting, 1852–1914):

$$\mathbf{S} = (1/\mu_0) \mathbf{E} \times \mathbf{B}. \quad (4.19)$$

Not only does  $|\mathbf{S}|$  give the intensity of the wave, but as the reader can readily verify,  $\mathbf{S}$  “points” in the direction of the propagation. The three vectors  $\mathbf{E}, \mathbf{B}, \mathbf{S}$  conveniently characterize a plane EM wave.

Unlike mechanical waves, no medium is required for EM propagation. Going back to Fig. 4.3 and the derivation of (4.13) we see that what produces, or “drives”, the magnetic field are the sheets of displacement current density,  $\epsilon_0 \partial \mathbf{E} / \partial t$ , which we find arranged in planes parallel to the  $x$ - $y$  plane. (The reader may find it helpful to look at Example 4.1 at the end of the chapter.) What is the origin of the displacement current sheets? The answer is that the sheets of changing magnetic field, which also lie in planes parallel to the  $x$ - $y$  plane but with  $\mathbf{B}$  pointing at right angles to  $\mathbf{E}$ , in turn generate the  $\mathbf{E}$  fields, entirely in accordance with Faraday’s Law. Of course, since the wave carries energy, something has to start the disturbance, just as one has to wiggle the end of a stretched string to get propagation there. Oscillating electric charges, be they comprised of the oscillating material components of a hot body, or be they the currents set up in a radio antenna, start the disturbance.

The radiation we have studied is *plane-polarized*, the “plane” being the  $x$ - $z$  plane to which the electric field is confined. There is nothing special about the  $x$ - $z$  plane. In general, light is comprised of radiation polarized in every direction. Mathematically, to get completeness, we add to the mode we have considered a second mode, which has fields rotated through ninety degrees about the  $z$ -axis relative to the first, i.e., the  $\mathbf{E}$  field points in the  $y$ -direction, and the  $\mathbf{B}$  field in the  $-x$ -direction. Together these are the two modes of polarization propagating in the  $z$ -direction which can be combined to give any desired net polarization.

In Chap. 1 we discussed how any function defined in one dimension between + and – infinity can be represented by the complete set of states

$$\sin kx, \quad \cos kx; \quad k \geq 0.$$

One can generalize this representation to three dimensions by taking the “wave vector”  $\mathbf{k}$  as truly a vector with a direction in space, rather than just as a number. In three dimensions the complete set of states becomes

$$\sin \mathbf{k} \cdot \mathbf{r}, \quad \cos \mathbf{k} \cdot \mathbf{r}$$

$$\sin \mathbf{k} \cdot \mathbf{r} = \sin(k_1 x + k_2 y + k_3 z), \quad \infty > k_1, k_2, k_3 > 0$$

where  $k_1, k_2, k_3$  are the cartesian components of wave vector  $\mathbf{k}$ .  $\sin \mathbf{k} \cdot \mathbf{r}$  represents a plane wave propagating in the direction of  $\mathbf{k}$ , see Problem 4.6.

To find a complete set of functions with which to represent any electromagnetic signal (natural motion), we need to similarly generalize the plane EM waves we have been studying. We enlarge our set of waves by allowing them to propagate in arbitrary directions, rather than just in the  $z$ -direction. Since we have been talking of traveling rather than standing waves, we also allow for negative values of  $\mathbf{k}$ . A complete set of EM waves is accordingly

$$\mathbf{E} = \mathbf{E}_0 \sin(\mathbf{k} \cdot \mathbf{r} - \omega t), \quad \text{with} \quad (4.20a)$$

$$\mathbf{B} = \mathbf{B}_0 \sin(\mathbf{k} \cdot \mathbf{r} - \omega t), \quad \infty > k_1, k_2, k_3 > -\infty. \quad (4.20b)$$

Since the set is again to be transverse, we must further restrict  $\mathbf{E}_0$  to take up an orientation perpendicular to the wave vector  $\mathbf{k}$ , and  $\mathbf{B}_0$  will be perpendicular to  $\mathbf{E}_0$ . Once such an orientation has been chosen (determination of plane of polarization), the complete set of modes is obtained by allowing, for each choice of  $\mathbf{k}$ , one further choice of  $\mathbf{E}_0$ , that perpendicular to the original choice. With two polarizations chosen for every  $\mathbf{k}$  in (4.20a,b), we have the desired complete set of functions for expressing the propagation of any EM signal in free space.

We recall from Chap. 3 (Problem 3.14) that in a one dimensional situation where reflections can be neglected, a time harmonic signal will be carried by a single traveling wave. Much the same considerations apply in analyzing the radiation from, say, a distant television antenna. In the vicinity of the antenna, i.e., only a small number of wavelengths or less away, the radiation pattern will be quite complicated. However, many wavelengths away it will be spherical, and radiation of each frequency will be carried by a single traveling wave, in this case spherical. Such a spherical wave can in turn be resolved into plane waves of the type in (4.20a,b), since these constitute a complete set. In general, the dimensions of the receiving antenna are small indeed compared to the radius of curvature of the transmitting wave, which means that the signal of a given frequency received by the antenna can for practical purposes be described by one wave of the set (4.20a,b), that heading in the direction of the receiver. Similarly the spherical traveling waves carrying the sun's radiation will converge at the usual focal point of a simple lens, indicating that the lens, tiny on a solar scale, is receiving plane waves.

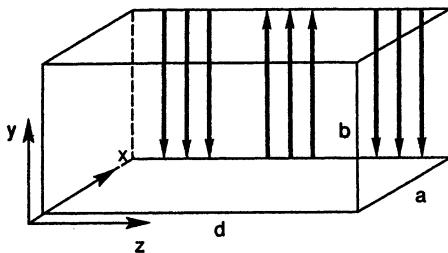
In the next section we shall be concerned with electromagnetic radiation, again in source-free space, but confined within boundaries, i.e., radiation in cavities. We shall see that the solutions obtained there have much in common with those obtained here for unconfined free space, indeed, they are combinations of a finite subset.

## 4.4 Distributed Electromagnetic Systems – Cavities

In Chap. 2 we analyzed the dynamics of mechanical and electrical systems with one degree of freedom. In Chap. 3 we were concerned with distributed mechanical systems. We must now study distributed electrical systems. We have seen that when one wishes to construct electrical oscillators one employs as the timing element  $L\text{-}C$  lumped circuits. These display the same characteristic resonance response as their mechanical counterparts, springs and pendula. However, as the resonance frequency increases to beyond the usual radio-frequency range the lumped  $L\text{-}C$  circuits become unsatisfactory due both to losses by radiation, and also to physical difficulties of making sufficiently precise circuit elements with very small values of  $L$  and  $C$ . Unwanted capacitance between the conductors comprising  $L$ , and similarly “stray” inductance become too large. The small elements automatically evolve into physical systems that are more accurately viewed as distributed, see Feynman’s *Lectures on Physics II* listed under Further Reading. The new systems are nothing but enclosures, cans or boxes of some shape, with walls constructed of materials that are good electrical conductors. We call these systems resonant cavities. To obtain high frequency oscillators one can couple the cavities into the remaining circuit with small wire antennas, or one can even couple them more directly to the current by having a free electron stream pass through them, as is done in the klystron.

Distributed electrical systems extend, of course, beyond closed systems (cavities) to open systems employed in the transfer of energy. At high frequencies the transfer of energy along ordinary wires is again accompanied by excess losses due to radiation, and one resorts to coaxial cables and waveguides. The analysis of these open systems depends upon our ensuing analysis of cavities. We restrict our rather detailed study in the text to just the cavities, which will be taken to be simple rectangular ones, with examples of open systems given in Problems.

Thus, we wish to develop the space–time behavior of the electric and magnetic fields in a rectangular cavity, measuring  $a \times b \times d$  (Fig. 4.5). No energy is added to or taken away from the fields confined to the interior of the metal box. We are, therefore, talking about the *natural* behavior of the fields. As in previous chapters one can subsequently discuss the *driven* behavior, i.e., what happens if a little antenna is placed inside the box and connected to an externally generated signal. (The details of a driven situation are treated in the Problems.) Although we can easily arrange not to supply power to the cavity beyond the first impulse that creates the fields, there will be, just as in the systems considered in earlier chapters, some losses, again basically of a frictional origin. These will slowly dampen whatever electromagnetic activity we have excited within the cavity. The losses will arise because our fields are accompanied by electric currents in the walls of the cavity, which are never perfect conductors. We are neglecting these losses.



**Fig. 4.5.** Rectangular cavity – a closed box with highly conductive walls. Also shown by the arrows, currents in the back plate for the TE<sub>103</sub> mode

The possible electric and magnetic fields in the cavity are determined, much as was the displacement of the stretched string, by finding which fields are both solutions of the wave equations that govern them, and satisfy boundary conditions imposed by what happens at the walls of the cavity. We shall also have to think about the subsidiary conditions,  $\nabla \cdot \mathbf{E} = 0, \nabla \cdot \mathbf{B} = 0$ . The boundary conditions for the stretched string that we considered were simple, neither end could move. The boundary conditions for the electric and magnetic fields are less obvious. We have decided to work with cavity walls that are perfect conductors. In this case the electric fields will adjust themselves so that their tangential components vanish at the boundaries. Were this not the case, the walls would be carrying infinite currents, giving rise to infinite fields. In practice, with good but not perfect conductors, a very small tangential component of the field remains. Our idealization is giving us a practical but not perfect solution. Since the electric fields do not penetrate the walls, we might expect to find restrictions on the accompanying magnetic field. The actual condition turns out to be that the normal component of  $\mathbf{B}$  vanishes at the boundary. We have

$$\text{at boundary } \begin{cases} \mathbf{E}_{\parallel} = 0 \\ \mathbf{B}_{\perp} = 0 \end{cases}. \quad (4.21)$$

The second condition also requires (4.11),  $\nabla \cdot \mathbf{B} = 0$ , for its derivation, see Problems. In fact, for perfect conductors we can find the fields without the second condition. It is satisfied automatically if the boundary condition for the  $\mathbf{E}$  field is satisfied. Ultimately the reason for this is the internal consistency of the Maxwell equations.

In (4.20) we have given a complete set of functions for expressing the electromagnetic field in free space, i.e., no sources, and with boundaries at infinity. These functions satisfy the time-dependent Maxwell equations. It turns out that a suitable subset of these functions can likewise be made to satisfy the boundary conditions for our rectangular cavity, and to form a complete set, a basis, within the restricted space of the cavity.

The functions in (4.20) are traveling waves. To accommodate the given fixed boundaries of the cavity they must first be recombined to form standing waves. We shall show that the standing waves which form the basis in our cavity are

$$E_x = E_1 \cos \omega t \cos k_1 x \sin k_2 y \sin k_3 z , \quad (4.22a)$$

$$E_y = E_2 \cos \omega t \sin k_1 x \cos k_2 y \sin k_3 z , \quad (4.22b)$$

$$E_z = E_3 \cos \omega t \sin k_1 x \sin k_2 y \cos k_3 z . \quad (4.22c)$$

where  $E_1, E_2, E_3$ , are amplitudes,  $k_1, k_2, k_3$ , are wave vector components.

Let us look at (4.22a) in some detail. We have chosen sine functions for the  $y$  and  $z$  dependence of  $E_x$  so as to secure vanishing of the tangential component of  $\mathbf{E}$  at the boundaries of the cavity at  $y = 0$  and  $z = 0$ . Further, looking at all three equations, the boundary conditions restrict the components of the wave vector  $\mathbf{k}$  to the subset

$$k_1 = l\pi/a, \quad k_2 = m\pi/b, \quad k_3 = n\pi/d , \quad (4.23)$$

where  $l, m, n$  are the positive integers,  $1, 2, 3 \dots$ , and at most only one of them may vanish, and  $a, b, d$  are the stated dimensions of the cavity.

As previously,  $k^2 = \omega^2/c^2$ , where  $c$  is the speed of light. In consequence, the angular frequencies for natural motion are

$$\omega_{l,m,n}^2 = c^2(k_1^2 + k_2^2 + k_3^2), \quad (4.24a)$$

$$= c^2(l^2\pi^2/a^2 + m^2\pi^2/b^2 + n^2\pi^2/d^2) . \quad (4.24b)$$

We note that the phase in (4.22a) has been arbitrarily set equal to zero. Actually it would be determined by the initial conditions resulting from excitation of the cavity.

As in free space the condition  $\nabla \cdot \mathbf{E} = 0$  will be satisfied if  $\mathbf{k}$  is chosen perpendicular to  $\mathbf{E}$ . Here, this gives us the restriction

$$\mathbf{E} \cdot \mathbf{k} = E_1 k_1 + E_2 k_2 + E_3 k_3 = 0 . \quad (4.25)$$

Further,  $\nabla \cdot \mathbf{E} = 0$  accounts for our having selected the  $x$ -dependence of  $E_x$ , the  $y$ -dependence of  $E_y$ , etc. to be cosinusoidal.

So far we have not bothered about  $\mathbf{B}$ . It can be calculated from

$$-\partial \mathbf{B} / \partial t = \nabla \times \mathbf{E} \quad (4.10)$$

with the  $\mathbf{B}$  field depending on time as  $\sin \omega t$ . One again finds that  $\mathbf{B}$  is perpendicular to both  $\mathbf{E}$  and  $\mathbf{k}$ . In seeking all independent solutions for the cavity we may again allow the  $\mathbf{E}$  vectors to assume any two mutually perpendicular directions in the plane perpendicular to a given wave vector  $\mathbf{k}$ . The complete basis set is obtained by enumerating all possible wave vectors  $\mathbf{k}$ , i.e., all allowed choices of the trio  $l, m, n$  in (4.23), and then assigning both polarizations to each choice of  $\mathbf{k}$ .

The method of classification of modes we have just given is often used in formal developments of physical theories, for example, when working problems in thermodynamics and statistical mechanics. However, when working with more detailed, descriptive problems, one prefers a set of explicit modes that already satisfy the constraint  $\nabla \cdot \mathbf{E} = 0$ . Leaving out any time factor to stress the fact that we are dealing with a set of spatially complete modes, also suitable for driven problems, we have

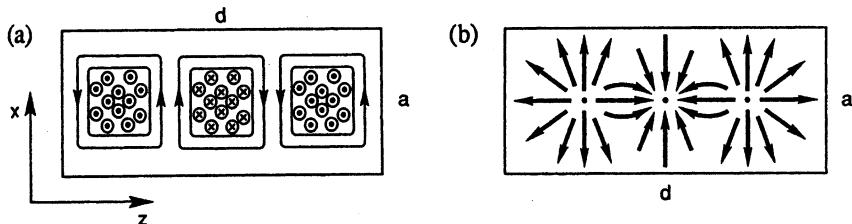
$$\begin{aligned}\mathbf{E}_{\text{TE}} = & -(k_2/k) \cos k_1 x \sin k_2 y \sin k_3 z \mathbf{i} \\ & + (k_1/k) \sin k_1 x \cos k_2 y \sin k_3 z \mathbf{j},\end{aligned}\quad (4.26a)$$

$$\begin{aligned}\mathbf{E}_{\text{TM}} = & [(k_1^2 + k_2^2)/k^2] \sin k_1 x \sin k_2 y \cos k_3 z \mathbf{k} \\ & - [k_1 k_3/k^2] \cos k_1 x \sin k_2 y \sin k_3 z \mathbf{i} \\ & - [k_3 k_2/k^2] \sin k_1 x \cos k_2 y \sin k_3 z \mathbf{j}.\end{aligned}\quad (4.26b)$$

Here the wave numbers  $k_1, k_2, k_3, k$  are the same as before, i.e., they satisfy (4.23–24a,b). Although this set is somewhat more specialized than the previous set, it favors the  $z$ -direction; it is complete. The modes as written are orthogonal. (To check, we take vector dot products and integrate over the volume of the cavity.) The normalization constants are chosen for convenience, see Problem 4.21. We note that the first kind of modes, those in (4.26a), are called “transverse electric” because we do not find any electric field in the  $z$ -direction. Similarly, the second set is called “transverse magnetic”, since for these modes the associated magnetic field in the  $z$ -direction, which we can get from (4.10), vanishes. These same modes can be used to describe the fields in waveguides for propagation in the  $z$ -direction. EM cavities have many applications in technology, for example, they form an integral component of lasers.

In Fig. 4.6 we illustrate the  $\text{TE}_{103}$  mode, viz.  $l = 1, m = 0, n = 3$ . The diagram corresponds to the instant  $\omega t = \frac{3\pi}{2}$  with the time dependent factor  $\cos \omega t$  multiplying  $\mathbf{E}(x, z)$ . From (4.26a) we see that the electric field will be given by

$$\begin{aligned}\mathbf{E}_{\text{TE}_{103}} &= k_1(k_1^2 + k_3^2)^{-1/2} \sin k_1 x \sin k_3 z \mathbf{j} \\ &= d(d^2 + 9a^2)^{-1/2} \sin \frac{\pi x}{a} \sin \frac{3\pi z}{d} \mathbf{j}.\end{aligned}$$



**Fig. 4.6.** (a) The  $\text{TE}_{103}$  mode. The closed lines represent the  $B$  field, the circles the displacement current,  $\varepsilon_0 \partial \mathbf{E} / \partial t$ . There is no variation in the  $y$ -direction. (b) Currents in the top plate. Lines of  $\varepsilon_0 \partial \mathbf{E} / \partial t$  in the cavity join to lines of  $\mathbf{J}$  in the plate, and the  $\mathbf{J}$  lines continue down the sides (Fig. 4.5);  $\nabla \cdot (\mathbf{J} + \varepsilon_0 \partial \mathbf{E} / \partial t) = 0$

In the diagram we illustrate the displacement current  $\varepsilon_0 \partial \mathbf{E} / \partial t$  by dots where it is pointing vertically upward, by crosses where it is pointing down. The continuous lines are the  $\mathbf{B}$  lines, calculated from (4.10). Once the magnetic fields are known we can also estimate the surface currents in the cavity walls from the boundary conditions. Using the reasoning of Example 4.1, we find that the cavity walls will carry surface currents per unit length,

$$J_s = B_{\tan,s} / \mu_0 , \quad (4.27a)$$

$B_{\tan,s}$  is the (tangential) magnetic field at the surface in question. At the inner surface of the cavity we have the vector relation

$$\mathbf{J}_s = \mathbf{n} \times \mathbf{B}_s / \mu_0 , \quad (4.27b)$$

where  $\mathbf{n}$  is a unit vector, pointing into the cavity in a direction normal to the surface, and  $\mathbf{B}_s$  is the magnetic field at the surface.

The surface currents are good estimates of actual currents even when the electric resistance of the cavity walls is finite. They then enable us to calculate a quantity  $W$ , namely, the average energy dissipated as joule heat during each cycle of operation. In analogy with  $L-C$  circuits we can define a quality indicator  $Q$  for a cavity oscillating in a given mode as

$$Q = 2\pi \frac{\text{time-average of energy stored in cavity}}{\text{energy loss per cycle of operation}} = 2\pi \frac{U}{W} . \quad (4.28a)$$

$U$ , the average energy stored in a given mode, is found by taking the volume integral of the energy density. The electric and magnetic energy densities can be seen from the previous section to be proportional to the square of the respective fields. Consequently, if  $\alpha$  is the damping constant for the fields, analogous to the constant in the lumped systems of Chap. 2, the energy will decay as  $e^{-2\alpha t}$ . We can write

$$U = U_0 e^{-2\alpha t} \quad (4.28b)$$

$$\frac{dU}{dt} = -2\alpha U \equiv -W/T = -Wf \quad (4.28c)$$

where  $f$  is the frequency,  $T$  the period of oscillation of the mode, and by defining  $W$  as the *average* loss per cycle,  $W/T$  is the rate of energy loss. Finally,

$$\alpha = \frac{Wf}{2U} = \frac{2\pi}{Q} \cdot \frac{f}{2} = \frac{\omega}{2Q} , \quad (4.28d)$$

giving the same relation between the damping coefficient and the quality factor as for a lumped system. We note that whereas the latter is characterized by a single  $Q$ , a distributed system possesses a separate  $Q$  for every mode of oscillation. We recall that a high  $Q$ , i.e., low loss, implies a sharp resonance response, that is, good performance in oscillatory or discriminatory applications. This is as true for a cavity “resonating” in one of its modes, as it is for an  $L-C$  circuit.

## 4.5 The Vector Potential and Related Solutions of Maxwell's Equations

Before undertaking a discussion of the vector potential we note two well-known properties of vector operators. If we form a vector field by taking the curl of some given initial vector field, then that second field will have zero divergence. In symbols, for any vector field  $\mathbf{F}(x, y, z)$

$$\nabla \cdot (\nabla \times \mathbf{F}) = 0. \quad (4.29a)$$

Equation (4.29a) is a vector identity: the divergence of a curl is zero. (See Example 4.2.) A second identity is that the curl of a gradient is zero:

$$\nabla \times (\nabla \psi) = 0 \quad (4.29b)$$

with  $\psi(x, y, z)$  being any scalar field.

Underlying our understanding of vector equations such as those of Maxwell is an important theorem, Helmholtz's theorem, which we will state without proof. The final form can be made plausible by reference to elementary electrostatics, as we do in one of the problems. The reader is referred for a full proof to the more advanced text by *Panofsky and Phillips* listed at the end of the chapter. The theorem states that any vector field  $\mathbf{F}(x, y, z)$  is uniquely determined if its curl and divergence are given, providing the source fields, i.e., its curl and divergence vanish at infinity. Suppose we have

$$\nabla \times \mathbf{F}(x, y, z) = \mathbf{c}(x, y, z) \quad (4.30a)$$

and

$$\nabla \cdot \mathbf{F}(x, y, z) = d(x, y, z) \quad (4.30b)$$

where  $\mathbf{c}$  and  $d$ , a vector and a scalar field, are the sources. The field  $\mathbf{F}$  is then

$$\mathbf{F}(x, y, z) = -\nabla \psi(x, y, z) + \nabla \times \mathbf{G}(x, y, z), \quad (4.31a)$$

where

$$\psi(x, y, z) = 1/4\pi \int \frac{d(x_0, y_0, z_0) dx_0 dy_0 dz_0}{[(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2]^{1/2}}, \quad (4.31b)$$

and

$$\mathbf{G}(x, y, z) = 1/4\pi \int \frac{\mathbf{c}(x_0, y_0, z_0) dx_0 dy_0 dz_0}{[(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2]^{1/2}}. \quad (4.31c)$$

Moreover, for this form for  $\mathbf{G}$

$$\nabla \cdot \mathbf{G} = 0. \quad (4.31d)$$

[Other forms for  $\mathbf{G}(x, y, z)$  are also permitted, we can add any field with zero curl to  $\mathbf{G}(x, y, z)$ . However, in every case  $\nabla \times \mathbf{G}$  will be as obtained from the solution given here.  $\mathbf{F}(x, y, z)$ , which is unique, has been determined.]

The field  $\mathbf{F}(x, y, z)$  is the sum of  $-\nabla\psi(x, y, z)$  and  $\nabla \times \mathbf{G}$ , that is, a *longitudinal* part (zero curl), and a *transverse* part (zero divergence). The longitudinal part is determined solely by the divergence of  $\mathbf{F}$ , while the transverse part is determined only by the curl of  $\mathbf{F}$ .

The scalar potential of (4.31b) is calculated in much the same way as the electric potential in electrostatics, by taking an element of source density,  $d(x_0, y_0, z_0)dx_0, dy_0, dz_0$ , dividing it by the distance between the source density and the observer at  $x, y, z$ , and adding up such terms until all the existing source density has been accounted for. The vector potential  $\mathbf{G}(x, y, z)$  is computed in the same manner, except that the source  $\mathbf{c}$  is a vector.

Let us apply the Helmholtz theorem in the given form to (4.11, 12). Since the divergence of  $\mathbf{B}$  is always zero we can write at once

$$\mathbf{B} = \nabla \times \mathbf{A}. \quad (4.32a)$$

Moreover, in accordance with (4.31d)

$$\nabla \cdot \mathbf{A} = 0, \quad (4.32b)$$

and, from (4.31c) we would have

$$\mathbf{A}(\mathbf{r}, t) = \mu_0 \int \frac{[\mathbf{j}(\mathbf{r}_0, t) + \varepsilon_0 (\partial \mathbf{E}(\mathbf{r}_0, t)/\partial t)] dv_0}{4\pi |\mathbf{r} - \mathbf{r}_0|}. \quad (4.32c)$$

Here  $\mathbf{A}(\mathbf{r}, t)$  is the vector potential for the magnetic field at the point  $\mathbf{r}(x, y, z)$ , and  $dv_0$  is used as an abbreviation for  $dx_0 dy_0 dz_0$ . We recall that the various fields depend on time as well as space. Equation (4.32c) is not our final result for  $\mathbf{A}(\mathbf{r}, t)$  because the source  $(\mathbf{j} + \varepsilon_0 \partial \mathbf{E}/\partial t)$  still contains one of the unknown fields,  $\mathbf{E}$ . We shall have to do more than directly apply the Helmholtz theorem to fully solve our problem, except in static situations, where  $\partial \mathbf{E}/\partial t = 0$ .

Again, because of the coupling between the  $\mathbf{E}$  and  $\mathbf{B}$  fields in the Maxwell equations, we can not apply the Helmholtz theorem directly to finding  $\mathbf{E}$ . However, here we can use the theorem to find a convenient result, leading towards solving our problem of finding the fields in terms of the sources. Using (4.32a) we rewrite (4.10)

$$\nabla \times \mathbf{E} = -\partial/\partial t(\nabla \times \mathbf{A}),$$

or

$$\nabla \times (\mathbf{E} + \partial \mathbf{A}/\partial t) = 0. \quad (4.33a)$$

Going back to auss's law (4.9), for the divergence of  $\mathbf{E}$ , and noting from (4.32b) that  $\nabla \cdot \mathbf{A} = 0$ , we find

$$\nabla \cdot (\mathbf{E} + \partial \mathbf{A}/\partial t) = \rho/\varepsilon_0 \quad (4.33b)$$

with  $\rho$  as the charge density. Equations (4.33a,b) are, thus, giving us both the curl and divergence of  $(\mathbf{E} + \partial \mathbf{A}/\partial t)$ , allowing us to use the Helmholtz theorem (4.31) to write

$$\mathbf{E} + \partial \mathbf{A}/\partial t = -\nabla \phi$$

or

$$\mathbf{E} = -\nabla \phi - \partial \mathbf{A}/\partial t \quad (4.34)$$

where

$$\phi(\mathbf{r}, t) = \int \frac{\varrho(\mathbf{r}_0, t) dv_0}{4\pi\epsilon_0 |\mathbf{r} - \mathbf{r}_0|}. \quad (4.35)$$

The potential  $\phi$  is the same as that encountered in electrostatics.

We will now find a modified expression for  $\mathbf{A}(\mathbf{r}, t)$ , which like (4.35) contains the sources only. Let us rewrite (4.12) so that the unknown field is on the left and the sources on the right:

$$\nabla \times \nabla \times \mathbf{A} + \mu_0\epsilon_0 \frac{\partial^2 \mathbf{A}}{\partial t^2} = \mu_0 \mathbf{j} - \mu_0\epsilon_0 \nabla (\partial \phi / \partial t). \quad (4.36)$$

Here we have substituted  $\mathbf{E}$  and  $\mathbf{B}$  in terms of the potentials. Since  $\phi$  is already known from (4.35), we considered  $\mu_0\epsilon_0 \nabla \partial \phi / \partial t$  to be a source term. A well-known vector identity states that

$$\nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A}.$$

But  $\nabla \cdot \mathbf{A} = 0$ . If we replace  $\mu_0\epsilon_0$  by  $c^{-2}$ , (4.36) becomes

$$\nabla^2 \mathbf{A} - 1/c^2 \frac{\partial^2 \mathbf{A}}{\partial t^2} = -\mu_0 \mathbf{j}_T, \quad (4.37a)$$

where

$$\mathbf{j}_T = \mathbf{j} - \epsilon_0 \nabla (\partial \phi / \partial t) \quad (4.37b)$$

the subscript T indicating total.

The vector potential satisfies a wave equation with a driving term  $-\mu_0 \mathbf{j}_T$ .

We are already familiar with a wave equation, namely, that governing the displacement of a string, equation (3.2):

$$\frac{\partial^2 y}{\partial x^2} - 1/s^2 \frac{\partial^2 y}{\partial t^2} = F(x, t)/\varrho.$$

We remember that in a "source"-free region of the string, i.e., where the externally applied force  $F(x, t)$  vanishes, the solution can be written in the form [see (3.5)]

$$y(x, t) = Y(t \pm x/s)$$

where  $s$  is the wave speed.

Beyond that, the wave equation (4.37a) for  $\mathbf{A}(\mathbf{r}, t)$  resembles another equation whose solution is now known to us, the so-called Laplace equation for  $\phi$ : combining (4.9) and (4.34) and, again, recalling that  $\nabla \cdot \mathbf{A} = 0$ , we find

$$\nabla^2 \phi = -\varrho/\epsilon_0, \quad (4.38)$$

which has as its solution precisely (4.35).

These two solutions, (4.35) and (3.5), make plausible the actual solution of (4.37a) for  $\mathbf{A}(\mathbf{r}, t)$ :

$$\mathbf{A}(\mathbf{r}, t) = \mu_0 \int \frac{\mathbf{j}_T(r_0, t') dv_0}{4\pi |\mathbf{r} - \mathbf{r}_0|}, \quad (4.39a)$$

where

$$t' = t - |\mathbf{r} - \mathbf{r}_0| / c. \quad (4.39b)$$

This result for  $\mathbf{A}(\mathbf{r}, t)$  is referred to as a retarded solution because the current density  $\mathbf{j}_T$  is evaluated not at the time  $t$ , but at the earlier time  $t - |\mathbf{r} - \mathbf{r}_0| / c$ , i.e., at the time the field was “generated”. Just as for the waves on a string, information is being passed along at a fixed speed, but here in three dimensions with a diminishing amplitude. The equally valid mathematical solution that has  $t' = t + |\mathbf{r} - \mathbf{r}_0| / c$  has found little practical value to date. In the last section we show using the Green’s function technique that, indeed, the expression in (4.39a) is the physical solution of the wave equation (4.37a) for  $\mathbf{A}(\mathbf{r}, t)$ . There we regard the solution, in analogy with electrostatics, as a superposition of elemental solutions

$$d\mathbf{A}(\mathbf{r}, t) = \mu_0 \mathbf{j}_T(r_0, t') dv_0 / (4\pi |\mathbf{r} - \mathbf{r}_0|) \quad (4.39c)$$

corresponding to point sources  $\mathbf{j}_T(r_0, t') dv_0$ .

Equations (4.37–39a,b) for the potentials  $\phi$  and  $\mathbf{A}$  respectively, and (4.32, 34, 35) giving the fields  $\mathbf{B}$  and  $\mathbf{E}$  in terms of the potentials, complete the work of finding these fields in terms of the sources  $\varrho$  and  $\mathbf{j}$ , that is, of solving Maxwell’s equations. In Table 4.1 we summarize the equations that govern the electromagnetic fields, and their solutions.

The solution (4.39c) for  $\mathbf{A}(\mathbf{r}, t)$  (a Green’s function), is to be compared to the solution in Problem 3.16 for the disturbance caused by a point force on a long string, i.e., when we can neglect reflections from the ends of the string. It follows that we can use (4.39) to directly calculate the electromagnetic radiation due to any space and time source distribution (antenna), providing the radiation is into free space. In the next and last section of this chapter we illustrate this by finding the radiated fields of a simple “dipole” antenna.

A last remark is to point out again that our potentials  $\mathbf{A}$  and  $\phi$  are not unique, even though  $\mathbf{E}$  and  $\mathbf{B}$  are uniquely determined by the sources. Other choices (gauges) for  $\mathbf{A}$  and  $\phi$  are also possible. Our choice is called the “transverse gauge” by virtue of the relation  $\nabla \cdot \mathbf{A} = 0$ . It has its origin in our statement of the Helmholtz theorem, where we chose  $\mathbf{G}$  to give  $\nabla \cdot \mathbf{G} = 0$ . Because certain problems in modern physics are best done in this gauge it is also sometimes called the quantum mechanical gauge. In one of the problems we show how to change  $\phi$  when we convert  $\mathbf{A}$  to a different gauge, i.e., take  $\nabla \cdot \mathbf{A} \neq 0$ .

**Table 4.1.** Summary of basic relations for electromagnetic fields

Integral Form of Maxwell's Equations	Differential Form of Maxwell's Equations
$\oint \mathbf{E} \cdot d\mathbf{S} = \epsilon_0^{-1} \int \rho dv$	$\nabla \cdot \mathbf{E} = \rho/\epsilon_0$
$\oint \mathbf{E} \cdot d\mathbf{l} = - \int \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{S}$	$\nabla \times \mathbf{E} = -\partial \mathbf{B}/\partial t$
$\oint \mathbf{B} \cdot d\mathbf{S} = 0$	$\nabla \cdot \mathbf{B} = 0$
$\oint \mathbf{B} \cdot d\mathbf{l} = \mu_0 \int (\mathbf{j} + \epsilon_0 \partial \mathbf{E} / \partial t) \cdot d\mathbf{S}$	$\nabla \times \mathbf{B} = \mu_0 [\mathbf{j} + \epsilon_0 \partial \mathbf{E} / \partial t]$

In the transverse gauge,  $\nabla \cdot \mathbf{A} = 0$ , and

vector potential  $\mathbf{A}(\mathbf{r}, t) = \mu_0 \int \mathbf{j}_T(\mathbf{r}_0, t') / (4\pi|\mathbf{r} - \mathbf{r}_0|) dv_0$   
with  $\mathbf{j}_T = \mathbf{j} - \epsilon_0 \nabla(\partial \phi / \partial t)$

scalar potential  $\phi(\mathbf{r}, t) = \int \rho(\mathbf{r}_0, t) / (4\pi|\mathbf{r} - \mathbf{r}_0|) dv_0$   
with  $t' = t - |\mathbf{r} - \mathbf{r}_0|/c$ .

$\mathbf{E} = -\nabla\phi - \partial \mathbf{A} / \partial t \quad \mathbf{B} = \nabla \times \mathbf{A}$ .

$\nabla \equiv \mathbf{i} \frac{\partial}{\partial x} + \mathbf{j} \frac{\partial}{\partial y} + \mathbf{k} \frac{\partial}{\partial z}$ .  $\oint \nabla \cdot \mathbf{F} dv = \oint \mathbf{F} \cdot d\mathbf{S}$ .  
 $\oint (\nabla \times \mathbf{F}) \cdot d\mathbf{S} = \oint \mathbf{F} \cdot d\mathbf{l}$ .

## 4.6 Dipole Radiation

As an application of the retarded potentials we shall study radiation from a *dipole* antenna. The term dipole originates from so-called “*multipole*” expansions of fields in powers of the ratio of the dimensions of the antenna  $l$  to the distance from the antenna to the observer  $|\mathbf{r}|$ . We shall assume that this ratio is small  $|\mathbf{r}| \gg l$ . To obtain the conventional dipole radiation one must further take the dimensions of the antenna as being small compared to the wavelengths of the radiation, and these in turn small compared to  $r$ . Thus, we will be seeking the total net power radiated from a harmonically fed antenna of length  $l$ , with  $|\mathbf{r}| \gg \lambda \gg l$ . The result finds application in electrical communication and in atomic physics.

We start by splitting the retarded expression for the vector potential  $\mathbf{A}(\mathbf{r}, t)$ , (4.39a) into two parts:

$$\mathbf{A}(\mathbf{r}, t) = \mathbf{A}_1(\mathbf{r}, t) + \mathbf{A}_2(\mathbf{r}, t), \quad (4.40a)$$

$$\mathbf{A}_1(\mathbf{r}, t) \equiv \frac{\mu_0}{4\pi} \int \frac{\mathbf{j}(\mathbf{r}_0, t') dv_0}{|\mathbf{r} - \mathbf{r}_0|}, \quad (4.40b)$$

$$\mathbf{A}_2(\mathbf{r}, t) \equiv -\frac{1}{4\pi c^2} \int \frac{\nabla_{r_0} [\partial\phi(\mathbf{r}_0, t')/\partial t'] dv_0}{|\mathbf{r} - \mathbf{r}_0|}, \quad (4.40c)$$

with again  $t' = t - |\mathbf{r} - \mathbf{r}_0|/c$ ; in (4.40c)  $\nabla_{r_0}$  involves differentiation with respect to the  $r_0$  coordinates.

Now,

$$\nabla \times \mathbf{A}_2(\mathbf{r}, t) = 0. \quad (4.41)$$

We can easily verify this by noting that for any one component of  $\nabla$ , say the  $x$  component, we can write

$$\begin{aligned} \frac{\partial}{\partial x} \mathbf{i} \times \mathbf{A}_2 &= \frac{1}{4\pi c^2} \int \left( \frac{\partial}{\partial x_0} \frac{1}{|\mathbf{r} - \mathbf{r}_0|} \right) \mathbf{i} \times \nabla_{r_0} [\partial\phi(\mathbf{r}_0, t')/\partial t'] dv_0 \\ &= \frac{-1}{4\pi c^2} \int \frac{1}{|\mathbf{r} - \mathbf{r}_0|} \frac{\partial}{\partial x_0} \mathbf{i} \times \nabla_{r_0} (\partial\phi(\mathbf{r}_0, t')/\partial t') dv_0. \end{aligned}$$

Here in the first line we have taken  $\partial/\partial x$  under the integral sign and then changed the variables of differentiation. To get the second line we have done an integration by parts, noticing that at  $x_0 = \pm\infty$  the relevant integral over  $dy_0 dz_0$  only is a vanishing quantity. Taking all three components together, one sees that one is left with a curl of a gradient in the integrand, which vanishes. Since the curl of  $\mathbf{A}_2(\mathbf{r}, t)$  vanishes we can obtain the magnetic field  $\mathbf{B}(\mathbf{r}, t)$  from  $\mathbf{A}_1(\mathbf{r}, t)$  alone. Having found  $\mathbf{B}(\mathbf{r}, t)$  we shall employ Maxwell's equation (4.12) to find  $\mathbf{E}(\mathbf{r}, t)$  directly. In this way we can avoid having to find  $\mathbf{A}_2(\mathbf{r}, t)$  and  $\phi(\mathbf{r}_0, t)$  altogether.

If we now take the distance between the observer and the antenna to be much longer than the antenna dimensions, we may replace  $|\mathbf{r} - \mathbf{r}_0|^{-1}$  in (4.40b) by  $r^{-1}$  and factor it out of the integral, giving

$$\mathbf{A}_1(\mathbf{r}, t) \cong \frac{\mu_0}{4\pi r} \int \mathbf{j}(\mathbf{r}_0, t') dv_0. \quad (4.42)$$

The current density  $\mathbf{j}(\mathbf{r}, t)$  is in general given by

$$\mathbf{j}(\mathbf{r}, t) = \varrho(\mathbf{r}, t) \mathbf{u}(\mathbf{r}, t), \quad (4.43)$$

where  $\varrho$  is the charge density, and  $\mathbf{u}$  the charge velocity. If we assume the charge to be divided into small aggregates of charge  $q_i$

$$q_i = \int \varrho dv_i$$

moving with velocity  $\mathbf{u}_i$ , then  $\int \mathbf{j}(\mathbf{r}_0, t) dv_0$  just becomes  $\sum_i q_i \mathbf{u}_i$ . The overall charge distribution for a radio antenna or, for that matter, for any given atom will, in general, be neutral. We shall suppose that it is the positive charges that are moving with respect to equal and opposite negative charges which are fixed.

Defining dipole moments by

$$\mathbf{p}_i(t) \equiv q_i \mathbf{r}_i(t)$$

with  $\mathbf{r}_i$  the location of the  $i$ th charge, and differentiating with respect to time, we have  $\dot{\mathbf{p}}_i(t) = q_i \mathbf{u}_i$ . Restricting our summation to the mobile positive charges, we have

$$\sum_i q_i \mathbf{u}_i = \sum_i \dot{\mathbf{p}}_i(t)$$

i.e., time derivatives of the dipole moments.

Going back to our assumption that the antenna length  $l$  is small compared to the radiated wavelength  $2\pi l/\lambda \ll 1$ , we can neglect any difference in phases due to retardation originating from the finite extent of the antenna. We can then assume a simple harmonic time dependence for the total dipole moment of the antenna:

$$\mathbf{p}(t) = \sum_i \mathbf{p}_i(t) \equiv \mathbf{p}_0 e^{i\omega t}.$$

Using the definition of the retarded time  $t'$ ,

$$\omega t' = \omega t - \left(\frac{\omega}{c}\right) r,$$

we finally have

$$\int \mathbf{j}(\mathbf{r}_0, t') dv_0 = i\omega \mathbf{p}_0 e^{i\omega t'} = i\omega \mathbf{p}(t) e^{-ikr} = e^{-ikr} \frac{d\mathbf{p}(t)}{dt},$$

with  $k = \omega/c$ . Equation (4.42) now gives

$$\mathbf{A}_1(\mathbf{r}, t) = \frac{\mu_0}{4\pi r} e^{-ikr} \frac{d\mathbf{p}(t)}{dt}. \quad (4.44)$$

The calculation of  $\nabla \times \mathbf{A}_1(\mathbf{r}, t) = \mathbf{B}(\mathbf{r}, t)$  and  $\nabla \times \mathbf{B}(\mathbf{r}, t) = \epsilon_0 \mu_0 \partial \mathbf{E} / \partial t$ , the latter equation being Maxwell's equation (4.12) in free space, is now entirely straightforward, and is posed in one of the Problems. Further, by integrating the Poynting vector  $\mathbf{S}$ ,

$$\mathbf{S} = \frac{1}{\mu_0} \mathbf{E} \times \mathbf{B} \quad (4.19)$$

over the surface of a sphere of radius  $r$ , we find in the problems that the total power crossing the surface of the sphere is independent of  $r$ , since by energy conservation it should be in steady-state, and is given by the well-known formula

$$\bar{P} \equiv \int \bar{S}(r) r^2 d\Omega = p_0^2 \omega^4 / (12\pi\epsilon_0 c^3). \quad (4.45)$$

The bars over  $P$  and  $S$  indicate that a time average over each cycle has been taken.

## 4.7 Electromagnetism and the Green's Function

In Sect. 4.5 plausibility arguments were given for finding the solution to the wave equation for the driven vector potential  $A(\mathbf{r}, t)$ :

$$\nabla^2 \mathbf{A} - 1/c^2 \frac{\partial^2 \mathbf{A}}{\partial t^2} = -\mu_0 \mathbf{j}_r \quad (4.37a)$$

$$\mathbf{j}_r = \mathbf{j} - \epsilon_0 \nabla (\partial \phi / \partial t) \quad (4.37b)$$

Thinking back to previous sections on the Green's function, 2.4, 3.5, the reader will correctly infer that the Green's function technique will again provide a systematic solution. Before treating (4.37a,b) we will consider a related but somewhat simpler problem, with a familiar solution.

The scalar potential of a static point charge of magnitude  $q$  located at  $\mathbf{r}_0$  is given by the expression:

$$\phi_q(\mathbf{r}|\mathbf{r}_0) = \frac{q}{4\pi\epsilon_0 |\mathbf{r} - \mathbf{r}_0|} \quad (4.46)$$

We now assert that for a charge of magnitude unity the expression (4.46) is the Green's function suitable for problems in electrostatics, providing the boundary can be taken at infinity:

$$G(\mathbf{r}|\mathbf{r}_0) = \frac{1}{4\pi\epsilon_0 |\mathbf{r} - \mathbf{r}_0|} \quad (4.47)$$

The electrostatic potential  $\phi(\mathbf{r})$  in general satisfies the Poisson equation (after the French mathematician S.D. Poisson, 1781–1840):

$$\nabla^2 \phi(\mathbf{r}) = -\rho(\mathbf{r})/\epsilon_0 \quad (4.38)$$

where  $\rho(\mathbf{r})$  is the charge density (source of the field). Accordingly, the appropriate Green's function  $G(\mathbf{r}|\mathbf{r}_0)$  is defined by the equation:

$$\nabla^2 G(\mathbf{r}|\mathbf{r}_0) = -\delta(\mathbf{r} - \mathbf{r}_0)/\epsilon_0 \quad (4.48)$$

In addition, since we are interested in problems where the boundary is at infinity, we need to specify  $G$  there. If we choose  $G$  in this way, we will be able to use the magic rule (Sects. 3.5 and 5.5) to solve (4.38) for the potential  $\phi(\mathbf{r})$ .

It is easy to show that the Green's function in (4.47) satisfies (4.48) when  $\mathbf{r} = \mathbf{r}_0$ . (The reader could, for example, verify that  $\partial^2 / \partial x^2 R$  equals  $-R^{-3}[1 - 3(x - x_0)^2/R^2]$  with  $R^2$  given by  $[(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2]$ .

We can show that at the source,  $G$  satisfies the governing equation (4.48) by integrating the equation over the region of the source. This is the same procedure as was followed in Chaps. 2 and 3 for the harmonic oscillator and stretched string, respectively. In this way the influence of the source is replaced by an additional boundary condition on the Green's function. We now construct a vanishingly small sphere of radius  $r_s$  about the source point at  $\mathbf{r}_0$ , and a large concentric sphere of radius  $r_l$ . Then,  $G$  will be defined in the region bounded by the two spheres by (4.48), which has a vanishing right side within the region, along with the conditions at the bounding surfaces.

Integrating the Poisson equation (4.48) over the inner sphere we get:

$$\int \int \int_{|\mathbf{r}-\mathbf{r}_0|=0}^{|\mathbf{r}-\mathbf{r}_0|=r_s} \nabla^2 G(\mathbf{r}|\mathbf{r}_0) d\mathbf{v}_0 = -\frac{1}{\epsilon_0} \int \int \int_{|\mathbf{r}-\mathbf{r}_0|=0}^{|\mathbf{r}-\mathbf{r}_0|=r_s} \delta(\mathbf{r}|\mathbf{r}_0) d\mathbf{v}_0 \quad (4.49a)$$

$$\text{or: } \int \nabla_r G(\mathbf{r}|\mathbf{r}_0) d\mathbf{S} = -\frac{1}{\epsilon_0} \quad (4.49b)$$

Here we have used Gauss's theorem to reduce the left side of (4.49a), ( $\nabla^2$  is **div.grad**). As we know from electrostatics the boundary condition described by (4.49b) is satisfied by the expression  $1/(4\pi\epsilon_0|\mathbf{r}-\mathbf{r}_0|)$  in (4.46). The gradient of  $G(\mathbf{r}|\mathbf{r}_0)$  yields the electrostatic field for a point charge as given by Coulomb's law:  $|\mathbf{E}| = 1/(4\pi\epsilon_0|\mathbf{r}-\mathbf{r}_0|^2)$ . Writing for  $d\mathbf{S}$ , an element of surface area of the small sphere,  $|\mathbf{r}-\mathbf{r}_0|^2 d\Omega$ ,  $d\Omega$  an element of solid angle, we easily see that (4.49b) is satisfied. Alternatively, we can write (4.49b) as:

$$\int \mathbf{E}(\mathbf{r}|\mathbf{r}_0) d\mathbf{S} = -\frac{1}{\epsilon_0} \quad (4.50)$$

Obviously, from Gauss's law, (4.50) holds. It follows from either argument that  $G(\mathbf{r}|\mathbf{r}_0)$  is a solution of the defining wave equation (4.48) even arbitrarily close to the singularity centered at  $\mathbf{r} = \mathbf{r}_0$ .

Since we are interested in electrostatic problems that specify a vanishing potential  $\phi$  at infinity, we know from the magic rule (see Sect. 3.5, or the three dimensional formulation given in Sect. 5.5) that  $G$  must vanish there as well. We have:

$$G(\mathbf{r}_l|\mathbf{r}_0) \rightarrow 0 \quad \text{as } |\mathbf{r}_l| \rightarrow \infty \quad (4.51)$$

Clearly,  $G$  as given in (4.47) satisfies this second requirement as well as the differential equation, (4.48).

With the magic rule we now obtain the standard result of electrostatics for:

$$\phi(\mathbf{r}) = \int G(\mathbf{r}|\mathbf{r}_0) \varrho(\mathbf{r}_0) d\mathbf{v}_0 \quad (4.52a)$$

$$= \frac{1}{4\pi\epsilon_0} \int \frac{\varrho(\mathbf{r}_0) d\mathbf{v}_0}{|\mathbf{r}-\mathbf{r}_0|} \quad (4.52b)$$

In special situations, such as problems with high symmetry, the Maxwell equations can be solved directly from their integral form. Electrostatics can

be built up directly from the known fields and potentials of individual point charges; each point charge has, after all, full spherical symmetry. In general, however, the Maxwell equations require rather far reaching mathematical methods for their formulation and systematic solution. When, in the course of solving these equations, we come to our present task of solving (4.37a), the wave equation for the vector potential  $\mathbf{A}(\mathbf{r}, t)$ , it is difficult to see how we can avoid the full mathematics.

To solve (4.37a) we need to modify the procedure we have just used to solve the Poisson equation owing to the addition of time dependence in the source, as well as in the wave equation itself. It will soon become evident that we need to weigh the source of the Green's function with the source density, solving:

$$\nabla^2 G_A(\mathbf{r}|\mathbf{r}_0|t) - 1/c^2 \partial^2 G_A(\mathbf{r}|\mathbf{r}_0|t)/\partial t^2 = -\mu_0 \mathbf{j}_r(\mathbf{r}_0, t) \delta^3(\mathbf{r} - \mathbf{r}_0) \quad (4.53)$$

We solve this vector equation by treating it as three independent equations, one equation for each cartesian component.

In Problem (4.20) the reader is guided through much the same steps as in the preceding development to show that the solution to (4.53) is given by:

$$\mathbf{G}_A(\mathbf{r}|\mathbf{r}_0|t) = \mu_0 \mathbf{j}_r(\mathbf{r}_0, t - |\mathbf{r} - \mathbf{r}_0|/c)/(4\pi|\mathbf{r} - \mathbf{r}_0|) \quad (4.54)$$

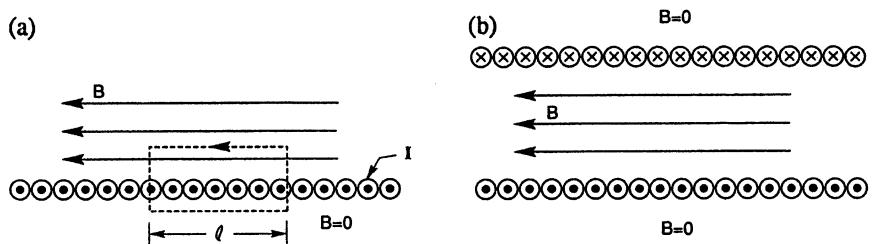
Finally, superposition (integrating (4.53) over  $d\mathbf{v}_0$ ) will yield:

$$\mathbf{A}(\mathbf{r}, t) = \mu_0 \int \frac{\mathbf{j}_r(\mathbf{r}_0, t')}{4\pi|\mathbf{r} - \mathbf{r}_0|} d\mathbf{v}_0 \quad (4.39a)$$

$$t' = t - |\mathbf{r} - \mathbf{r}_0|/c \quad (4.39b)$$

**Example 4.1.** A series of infinitely long straight wires is arranged so that they all lie side by side in an infinite plane, as shown in Fig. 4.7a. If each wire carries a current  $I$ , and there are  $N$  wires for each cross-sectional meter, find (a) the effective surface current density, (b) the magnetic field that exists above this conductor, given that no magnetic field exists below it.

*Hint:* Use (4.4) (the integral form).



**Fig. 4.7.** (a) Example 4.1. A series of infinitely long straight parallel wires.  
(b) Example 4.1. Two planes of wires, superimposed

*Solution:* The surface current  $J_s$ , i.e., the total current per unit cross-sectional area is given by  $J_s = NI$ . In the absence of a time-varying electric field, (4.4) reads

$$\oint \mathbf{B} \cdot d\mathbf{l} = \mu_0 \int \mathbf{J} \cdot d\mathbf{S}.$$

As the circuit for the integral on the left, we pick a rectangular contour piercing the plane of conductors, and perpendicular to them, as shown by the dashed line in Fig. 4.7a. By symmetry the magnetic field is horizontal, and there is no magnetic field in the region of the bottom contour. Integration along the top gives  $Bl$ , where  $l$  is the width of the rectangle. The current threading the area around which we have just integrated is obviously

$$\int J \cdot dS = J_s l = NI l.$$

Applying (4.4) to the above, we have

$$|B|l = \mu_0 NIl$$

$$B = \mu_0 NI.$$

The total magnetic field induced by the conductors will have the direction shown in Fig. 4.7a. From symmetry, one might expect equal and opposite fields above and below, each of magnitude  $\mu_0 NI/2$ . In fact there might also be a return sheet of current somewhere else, giving rise to the situation in Fig. 4.7b, in complete analogy to the electric capacitor (fields reinforcing each other inside, canceling outside).

This example illustrates the production of sheets of parallel magnetic fields by sheets of current, with the current flowing in a direction perpendicular to the induced fields. We find this situation in the plane EM waves of Sect. 4.3 where the current sheets are produced, not by charges flowing down wires, but by the displacement currents  $\epsilon_0 \partial \mathbf{E} / \partial t$ . We also find it in cavities, where surface currents  $J_s$  keep the magnetic fields from entering the conducting walls, confining them to the interior of the cavity.

**Example 4.2.** Show that the divergence of a curl field always vanishes.

*Solution:* The divergence of any vector  $\mathbf{F}$  is given by

$$\begin{aligned} \nabla \cdot \mathbf{F} &= \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z} \quad \text{and} \\ \nabla \cdot \nabla \times \mathbf{F} &= \nabla \cdot \left( \frac{\partial}{\partial x} \mathbf{i} + \frac{\partial}{\partial y} \mathbf{j} + \frac{\partial}{\partial z} \mathbf{k} \right) \times (F_x \mathbf{i} + F_y \mathbf{j} + F_z \mathbf{k}) \\ &= \frac{\partial}{\partial x} \left( \frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z} \right) + \frac{\partial}{\partial y} \left( \frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x} \right) \\ &\quad + \frac{\partial}{\partial z} \left( \frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right). \end{aligned}$$

Remembering that second order partial derivatives are independent of the order of differentiation, we see that all the terms cancel in pairs, proving the identity.

**Example 4.3.** Maxwell's first equation,  $\nabla \cdot \mathbf{E} = \rho/\epsilon_0$ , is well illustrated in the operation of a vacuum diode. Here electrons are emitted thermally in vanishing electric field and with vanishing velocity at the cathode. They are accelerated as they cross an evacuated gap of width  $d$  to the anode, held at fixed potential  $V$  relative to the cathode. A current implies a charge density  $\rho$  (which will be negative, owing to the negative charge  $e$  of the electron). If we assume infinite plane electrodes parallel to the  $y$ - $z$  plane then  $\rho$  will depend on  $x$  only:  $\rho(x)$ . This charge density, by the Maxwell equation will influence the electric field  $E(x)$ , which will be in the  $x$  direction and negative, pointing from anode to the cathode at  $x = 0$ . Thus,  $\rho(x)$  affects the motion of the electrons. In steady state, however, the current/unit area,  $J$  must be independent of  $x$  and the time  $t$  if  $\rho(x)$  is to be constant in time. The problem is to first find the dependence of the electrostatic potential  $\phi(x)$ , and then the dependence of  $J$  on the applied potential  $V$ , both in steady state.

*Solution:* The current density  $J$  is related to the charge/volume  $\rho$  through

$$J = -\rho(x)v(x)$$

with  $J$  and  $v(x)$ , the electron velocity, both positive. In steady state

$$E(x) = -d\phi(x)/dx \quad \text{with} \quad \phi(d) = V, \phi(0) = 0.$$

Since the electrons start from rest, and are assumed to influence each other only through the average electric field  $E(x)$ , conservation of energy yields:

$$\frac{1}{2}mv^2(x) + e\phi(x) = 0.$$

Lastly,  $\nabla \cdot \mathbf{E} = \rho/\epsilon_0$  becomes

$$\frac{d^2\phi(x)}{dx^2} = -\frac{\rho(x)}{\epsilon_0}.$$

If we now eliminate  $\rho(x)$  in favor of  $J$ , then with

$$v(x) = [2|e|\phi(x)/m]^{1/2}$$

we get

$$\frac{d^2\phi(x)}{dx^2} = \frac{J}{\epsilon_0} \sqrt{\frac{m}{2|e|}} \frac{1}{\phi^{1/2}(x)}$$

or

$$\frac{d^2\phi}{dx^2} = CJ\phi^{-1/2}(x), \quad C \equiv \frac{1}{\epsilon_0} \sqrt{\frac{m}{2|e|}}.$$

This homogeneous second-order equation can be transformed to a first order equation by assuming that  $\phi(x)$  and  $x(\phi)$  are known, and letting:

$$y(\phi) \equiv d\phi(x)/dx .$$

The second order equation becomes

$$\frac{dy(\phi)}{d\phi} \frac{d\phi(x)}{dx} = CJ\phi^{-1/2}(x), \text{ or}$$

$$y dy = CJd\phi/\phi^{1/2} .$$

Upon integrating we find:

$$y^2/2 = 2CJ\phi^{1/2} .$$

The constants of integration vanish, since at  $x = 0$ ,  $\phi = 0$ ; the electric field  $-y(\phi = 0)$  also vanishes.

Thus,

$$d\phi(x)/dx = 2[CJ]^{1/2}\phi^{1/4}(x) .$$

Integrating again, we find:

$$\phi(x) = (3/2)^{4/3}[CJ]^{2/3}x^{4/3} ,$$

giving us the desired expression for the static potential  $\phi(x)$ . Further, the current  $J$  is related to the potential drop across the diode  $\phi(d) = V$ , by

$$J = 4/(9Cd^2)V^{3/2} .$$

This distinctly non-ohmic result for a vacuum diode is known as the Child-Langmuir law. The reader will find further analysis of these results for the diode in Problem 4.A.1.

**Example 4.4.** Rewrite the TE cavity modes of (4.26a) to describe em fields in a wave guide driven at  $z = 0$ , which are propagating in the positive  $z$  direction. Find the value of  $k_3$  corresponding to given values of  $l$  and  $m$ , and a given driving frequency  $\omega$ . Check that  $\nabla \cdot \mathbf{E} = 0$ .

*Hint:* See Problem 3.14

*Solution:* In analogy with Problem 3.14, the time dependence will be associated with traveling waves in the positive  $z$  direction,  $\cos(k_3 z - \omega t + \delta_{l,m})$ , where the propagation vector  $k_3$ , and the phase  $\delta_{l,m}$  depend on the mode, and  $\delta_{l,m}$  is determined by the boundary condition at the driven end,  $z = 0$ . With this time dependence we can satisfy both the homogeneous wave equation for  $\mathbf{E}(x, y, z; t)$  inside the guide, as well as the boundary condition as  $z \rightarrow \infty$  — that there is no signal traveling (back) in the negative  $z$  direction.

For the electric field in the mode  $l, m$  we write:

$$\begin{aligned}\mathbf{E}_{\text{TE}_{l,m}}(x, y, z; t) = & [-(k_m/k) \cos(k_l x) \sin(k_m y) \mathbf{i} \\ & +(k_l/k) \sin(k_l x) \cos(k_m y) \mathbf{j}] \cos(k_3 - \omega t + \delta_{l,m})\end{aligned}$$

where

$$k_l = l\pi/a, \quad k_m = m\pi/b, \quad l = 1, 2, 3 \dots, \quad m = 1, 2, 3 \dots.$$

To find  $k_3$  we substitute the given electric field into the homogeneous wave equation in free space for  $\mathbf{E}$  (see Problem 4.13):

$$\nabla^2 \mathbf{E} - \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0.$$

This equation must be satisfied by each vector component of  $\mathbf{E}$  individually. Adopting a short-hard notation, we write:

$$\mathbf{E}_{\text{TE}_{l,m}}(x, y, z; t) \equiv [E_1(x, y) \mathbf{i} + E_2(x, y) \mathbf{j}] \cos(k_3 z - \omega t + \delta_{l,m}).$$

We find, upon carrying out the indicated differentiation, that:

$$\begin{aligned}\nabla^2 E_1 &= \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) E_1 = -(k_l^2 + k_m^2) E_1, \\ \nabla^2 E_2 &= -(k_l^2 + k_m^2) E_2\end{aligned}$$

and, so

$$\nabla^2 \mathbf{E}_{\text{TE}_{l,m}} = -(k_l^2 + k_m^2 + k_3^2) \mathbf{E}_{\text{TE}_{l,m}}.$$

Carrying out the differentiations with respect to  $t$  as well, we find that the wave equation is satisfied if:

$$k^2 \equiv k_l^2 + k_m^2 + k_3^2 = \omega^2/c^2$$

relating  $k_3$  to the driving frequency  $\omega$ , and to  $k_l$  and  $k_m$ . See Problem 4.13 for further discussion.

Turning to  $\nabla \cdot \mathbf{E}_{\text{TE}_{l,m}}$ , we have

$$\begin{aligned}& \left( \frac{\partial}{\partial x} \mathbf{i} + \frac{\partial}{\partial y} \mathbf{j} + \frac{\partial}{\partial z} \mathbf{k} \right) \cdot E_1(x, y) \mathbf{i} \cos(k_3 z - \omega t + \delta_{l,m}) \\ &= \frac{k_l k_m}{k} \sin k_l x \sin k_m y \cos(k_3 z - \omega t + \delta_{l,m}) \\ &= -\nabla \cdot [E_2(x, y) \mathbf{j}] \cos(k_3 z - \omega t + \delta_{l,m})\end{aligned}$$

Evidently, the phase  $\delta_{l,m}$  must be the same for both vector components of  $\mathbf{E}_{\text{TE}_{l,m}}$  if the divergence of  $\mathbf{E}$  is to vanish, as it must in free space inside the guide.

We note that the modes have been designed to satisfy boundary conditions at the walls of the guide. The work of finding the electric field inside the guide can be completed by determining the amplitude and phase of each TE and TM mode from the driving fields or sources at  $z = 0$ .

## Problems

**4.1** When at 1 cm from an insulating plate measuring  $5 \text{ m} \times 5 \text{ m}$  an observer measures an electric field of  $4 \times 10^4 \text{ V/m}$ , estimate what field he would measure 80 m from the plate, assuming a uniform charge distribution on the plate.

**4.2** Use (4.4) to find a formula for the magnetic field surrounding a long wire carrying a time independent current  $I$ .

**4.3** A long wire is connected in series to a parallel plate capacitor, as shown in Fig. 4.8. If the wire carries current  $I$ , show that

- a)  $I = CdV/dt$
- b)  $I_{\text{displ}} = CdV/dt$

where  $I_{\text{displ}}$  is the total displacement current through the capacitor. Why is the result  $I = I_{\text{displ}}$  necessary for consistency in the applications of Ampere's Law (4.4) to this situation?

*Hint:* Consider the current threading through a surface which passes between the capacitor plates rather than through a surface cut by the incoming wire.

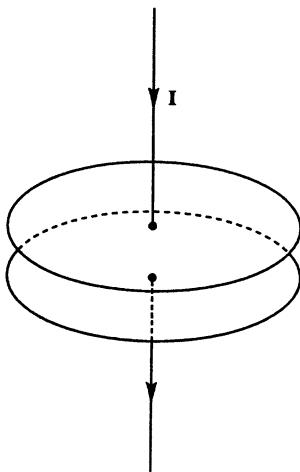


Fig. 4.8. Problem 4.3

**4.4** a) Show clearly how the physical interpretation of the divergence and the law of conservation of charge lead to the well-known relation

$$\nabla \cdot \mathbf{j} + \partial \varrho / \partial t = 0$$

where  $\mathbf{j}$  and  $\varrho$  are, respectively, the current and charge densities.

- b) Infer from (4.12) that  $\nabla \cdot (\mathbf{j} + \epsilon_0 \partial \mathbf{E} / \partial t)$  must equal 0.
- c) Use part (a) and another Maxwell equation to verify the result in (b).
- d) Show, assuming them to be uniformly distributed, the real and displacement currents on the plates and within the parallel plate capacitor of

Problem 4.3, as well as the magnetic field lines just without and within the plates. Discuss the application of (4.4) to this situation.

**4.5** It is well known that a particle with constant angular velocity  $\omega$  has a linear velocity  $\mathbf{v} = \omega \times \mathbf{r}$ .

a) Find  $\nabla \times \mathbf{v}$ .

(One may without loss of generality take  $\omega = |\omega| \mathbf{k}$ , with  $\mathbf{k}$  the unit vector in  $z$ -direction.)

b) Use your result in (a) to obtain  $\oint \mathbf{v} \cdot d\mathbf{r} / \int d\mathbf{S}$  where  $\int d\mathbf{S}$  is the surface area enclosed by the path  $\oint d\mathbf{r}$ , itself taken in the  $x$ - $y$  plane. Check by calculating  $\oint \mathbf{v} \cdot d\mathbf{r}$  directly for a circular path.

c) Write an analytic expression for the vector potential  $\mathbf{A}(\mathbf{r})$  giving rise to a uniform magnetic field  $\mathbf{B}$  in, say, the  $z$ -direction. Make a plot, showing lines of constant  $|\mathbf{A}(\mathbf{r})|$ . Calculate  $\nabla \cdot \mathbf{A}$ , and comment on the uniqueness of your result for  $\mathbf{A}(\mathbf{r})$ ; (see discussion of (4.31d)).

d) Distinguish between fluids flowing with velocities  $\mathbf{v}(\mathbf{r})$  that do or do not possess a curl. Does one of them possess an element of rotation?

**4.6** a) Consider the wave

$$\psi = \psi_0 e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)} = \psi_0 e^{i(k_x x + k_y y + k_z z - \omega t)} .$$

Show that points of constant phase lie on a plane.

b) Consider the longitudinal and transverse waves

$$\begin{aligned} \mathbf{F}_l &= F_0 e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)} \frac{\mathbf{k}}{|\mathbf{k}|} \\ \mathbf{F}_t &= F_0 e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)} \frac{\mathbf{k}_\perp}{|\mathbf{k}_\perp|} \quad \text{where} \quad \mathbf{k}_\perp \cdot \mathbf{k} \equiv 0 . \end{aligned}$$

Show that  $\nabla \cdot \mathbf{F}_t = 0$  and  $\nabla \times \mathbf{F}_l = 0$ .

Since any function, longitudinal or transverse, can respectively be written as a Fourier expansion of such waves, what can you conclude about the two alternative definitions of longitudinal and transverse mentioned in the text when there is a specified direction of propagation?

**4.7** a) Equation (4.28) shows  $\mathbf{E}$  and  $\mathbf{B}$  to be in phase for a transverse EM wave. Is this arbitrary?

b) Derive the relation  $E_0 = cB_0$  (4.17c).

**4.8** Use the vector identity valid for any two vector fields  $\mathbf{A}(x, y, z)$ ,  $\mathbf{B}(x, y, z)$ ,

$$\nabla \cdot (\mathbf{A} \times \mathbf{B}) = \mathbf{B} \cdot (\nabla \times \mathbf{A}) - \mathbf{A} \cdot (\nabla \times \mathbf{B})$$

to show that

$$\nabla \cdot \mathbf{S} \equiv \nabla \cdot (\mathbf{E} \times \mathbf{B} / \mu_0) = \frac{-\partial}{\partial t} \left[ \frac{B^2}{2\mu_0} + \frac{\epsilon_0 E^2}{2} \right] .$$

Explain why  $\mathbf{S}$ , the Poynting vector, can be interpreted as giving the energy flux.

**4.9** Sunlight strikes the earth, outside its atmosphere, with a time average intensity of 1400 watts/m<sup>2</sup>. Find  $E_0$ , the amplitude of the impinging electric field, assuming an incident EM wave which is plane.

**4.10** Use  $\nabla \cdot \mathbf{B} = 0$  to show that if the  $\mathbf{B}$  field does not enter a perfect conductor, its normal component must vanish at the surface.

*Hint:* Choose as the gaussian surface a little “pill box” with its sides intersected by the surface of the conductor.

**4.11** In a rectangular cavity the wave vectors  $k_1, k_2, k_3$  may be considered as forming a space, sometimes called reciprocal space. Show that the number of standing wave normal modes that may be associated with an element of volume  $\Delta^3 k$  in this space is

$$\Delta N = 2V/\pi^3 \Delta^3 k$$

where  $V$  is the volume of the cavity

$$\Delta^3 k = \Delta k_x \Delta k_y \Delta k_z .$$

**4.12** a) Calculate from the given electric field the magnetic field associated with the TE<sub>103</sub> mode, and check that it has been drawn in correctly in Fig. 4.6.

b) Calculate the current distribution associated with this mode which holds in the end wall where  $z = d$ .

**4.13** A TE mode

$$\mathbf{E} = E_0 \sin \frac{\pi x}{a} e^{i[\omega t - 2\pi z/\lambda_g]} \mathbf{j}$$

where  $\mathbf{j}$  is the unit vector in the  $y$ -direction, propagates down a rectangular waveguide of width  $\Delta x = a$ .

a) Show by using Maxwell's equations that inside the guide the free space wave equation is

$$\nabla^2 E - 1/c^2 \left( \frac{\partial^2 E}{\partial t^2} \right) = 0 .$$

b) Show that

$$\lambda_g = \lambda_0 / \sqrt{1 - (\lambda_0/2a)^2}$$

where  $\lambda_0 = 2\pi c/\omega$ .

c) Find  $\omega_c$ , the cut-off frequency, below which the mode will not propagate, i.e., where it becomes exponentially damped as a function of  $z$ .

d) Calculate  $\mathbf{B}$ , and draw an  $x$ - $z$  cross section, showing both the  $\partial \mathbf{E} / \partial t$  and  $\mathbf{B}$  fields in the guide over a distance  $\Delta z = \lambda_g$  at some instant of time. Take  $\omega > \omega_c$ .

*Hint:* see Example 4.4.

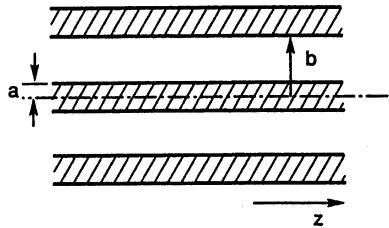


Fig. 4.9. Problem 4.14

**4.14** Power flows down a coaxial cable in its principal TEM mode. In the free space  $b > r > a$  (Fig. 4.9), the electric field  $\mathbf{E}$  is given by

$$\mathbf{E} = E_0 u(r) e^{i[\omega t - 2\pi z/\lambda]} \mathbf{a}_r$$

where  $r, z$  are cylindrical coordinates, and  $\mathbf{a}_r$  is the unit vector pointing in the radial direction.  $u(a) = 1$ .

a) On a side view such as the one shown in Fig. 4.9, illustrate with field lines the displacement current  $\epsilon_0 \partial \mathbf{E}_0 / \partial t$  over one wavelength of the transmitted signal. Find  $\partial r(z, t) / \partial t$ , the time rate of change of the charge per unit length on the inner conductor.

b) Similarly plot the magnetic fields. Which of Maxwell's equations are needed to find the field? Lastly, show also the current  $I_0 \cos(\omega t - 2\pi z/\lambda)$  in the conductors. Calculate  $I_0$ .

c) Use Gauss's law (4.1) to determine  $u(r)$ , noting that no charge density exists in the free space  $b > r > a$ . Find  $V_0(z, t)$ , the potential difference between the outer and the inner conductor. (The Lorentz force, form  $q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$  shows that no work is done by the magnetic field.)

d) With  $\mathbf{E}(r, z)$  and  $\mathbf{B}(r, z)$  in hand, check that all four of Maxwell's equations are satisfied providing  $2\pi/\lambda = \omega/c$ ,  $c$  the light velocity.

e) Find  $\mathbf{A}(r, z, t)$  and  $\phi(r, z, t)$  for  $b > r > a$ .

f) Calculate  $z_0 = V_0(z, t) / I_0(z, t)$ , the characteristic impedance of the cable. What happens qualitatively if some finite length of the cable is terminated by a circuit element whose impedance  $Z = z_0$ , what happens if  $Z \neq z_0$ ?

*Hint:* see Problem 6.11.

In cylindrical coordinates:

$$\nabla \cdot \mathbf{A} = 1/r \partial(r A_r) / \partial r + 1/r \partial A_\phi / \partial \phi + \partial A_z / \partial z$$

$$\nabla \Psi = \partial \Psi / \partial r \mathbf{a}_r + 1/r \partial \Psi / \partial \phi \mathbf{a}_\phi + \partial \Psi / \partial z \mathbf{a}_z ,$$

and see Problem 4.18 for the curl.

**4.15** a) Starting with (4.33a,b), show that adding a curl-free field  $-\nabla \chi$ , where  $\chi(\mathbf{r}, t)$  is any scalar field with appropriate boundary conditions to  $\mathbf{A}$ , not only leaves  $\mathbf{B}$  unchanged but  $\mathbf{E}$  as well, thereby guaranteeing unique solutions to Maxwell's equations.

*Hint:* Show that

$$\chi(\mathbf{r}, t) = -1/4\pi \int \frac{(\nabla')^2 \chi(\mathbf{r}') dv'}{|\mathbf{r} - \mathbf{r}'|}$$

before applying the Helmholtz vector theorem.

b) Show that in the Lorentz gauge  $\nabla \cdot \mathbf{A} = (-1/c^2)(\partial\phi/\partial t)$ , (4.31) becomes

$$\nabla^2 \mathbf{A} - \frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} = -\mu_0 \mathbf{j},$$

$$\nabla^2 \phi - \frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} = -\varrho/\epsilon_0.$$

What is  $\chi$ ?

**4.16** a) Show from (4.31a)

$$\mathbf{F}(x, y, z) = -\nabla\psi(x, y, z) + \nabla \times \mathbf{G}(x, y, z)$$

that for sources  $d$  and  $\mathbf{c}$ , (4.30a,b)

$$-\nabla^2 \psi = d \quad -\nabla^2 \mathbf{G} = \mathbf{c}$$

providing (4.31d),  $\nabla \cdot \mathbf{G} = 0$ , holds.

b) Accepting the “electrostatic” solution to  $-\nabla^2 \psi = d$ ,

$$\psi = \frac{1}{4\pi} \int \frac{d(r') dv'}{|\mathbf{r} - \mathbf{r}'|}$$

where  $\psi(\mathbf{r}) \rightarrow 0$  as  $|\mathbf{r}| \rightarrow \infty$ , can you find  $\mathbf{G}(\mathbf{r})$  satisfying the same boundary condition?

**4.17** Consult Table 4.1 to write down the solution for  $\mathbf{E}$  and  $\mathbf{B}$  in terms of the sources  $\varrho$  and  $\mathbf{j}$  when the latter do not depend on time. Under what circumstances can a vector field be written as the gradient of a scalar field only, as the curl of a vector field only, or a combination of both? Give reasons for your answers.

**4.18** Consider a straight wire extending from  $z' = -L/2$  to  $z' = +L/2$ , and carrying a time-independent current  $I$ .

a) Set up, but do not evaluate, an integral for the vector potential  $\mathbf{A}(r, z, \phi)$  in terms of a given current density  $\mathbf{j}(r', z', \phi')$ , and simplify to the thin wire at  $z = 0$ , carrying total current  $I$  from  $z' = -L/2$  to  $z' = +L/2$ .

b) Calculate  $\mathbf{B}(z, r, \phi) = \nabla \times \mathbf{A}$ , by differentiating the result in (a) under the remaining integral sign  $\int dz'$ , and then setting  $z = 0$ .

c) Complete the integration to find  $\mathbf{B}(r, 0, \phi)$ .

d) Consider your result for  $B$  as  $L \rightarrow \infty$ , and compare with the result obtained in Problem 4.2.

In cylindrical coordinates  $\nabla \times \mathbf{A} =$

$$\left( \frac{1}{r} \frac{\partial A_z}{\partial \phi} - \frac{\partial A_\phi}{\partial z} \right) \mathbf{a}_r + \left( \frac{\partial A_r}{\partial z} - \frac{\partial A_z}{\partial r} \right) \mathbf{a}_\phi + \frac{1}{r} \left[ \frac{\partial}{\partial r} (r A_\phi) - \frac{\partial A_r}{\partial \phi} \right] \mathbf{a}_z .$$

**4.19** From the static relation

$$d\mathbf{A} = \frac{\mu_0 I}{4\pi} \frac{dl}{|\mathbf{r} - \mathbf{r}'|}$$

derive the Biot-Savart law

$$d\mathbf{B} = \frac{\mu_0 I}{4\pi} \frac{dl \times (\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^3} .$$

*Hint:*  $\nabla \times (\psi \mathbf{F}) = \psi \nabla \times \mathbf{F} + (\nabla \psi) \times \mathbf{F}$ .

**4.20** a) Show that the vector potential  $\mathbf{G}_A(\mathbf{r}|0|t)$  induced by a point current density  $\mathbf{j}_T(0, t)\delta^3(\mathbf{r})$  at the origin satisfies the wave-equation

$$\nabla^2 \mathbf{G}_A - 1/c^2 \frac{\partial^2 \mathbf{G}_A}{\partial t^2} = 0$$

at every point  $\mathbf{r} \neq 0$ , if

$$\mathbf{G}_A(\mathbf{r}|0|t) = \mu_0 \mathbf{j}_T(0, t - r/c)/(4\pi r) .$$

*Hint:* In spherical coordinates

$$\nabla^2 \psi = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} \psi \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial}{\partial \theta} \psi \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 \psi}{\partial \phi^2} .$$

b) Derive a boundary condition for  $\mathbf{G}_A(\mathbf{r}|0|t)$  in analogy with (4.49b), to replace the effect of the drive in (4.53), here in the vicinity of  $\mathbf{r}_0 = 0$ .

c) Verify by direct calculation that  $\mathbf{G}_A(\mathbf{r}|0|t) = \mu_0 \mathbf{j}_T(0, t - r/c)/(4\pi r)$  satisfies the boundary condition derived in (b). In spherical coordinates

$$\nabla_r \equiv \frac{\partial}{\partial r} \mathbf{a}_r + \frac{1}{r} \frac{\partial}{\partial \Theta} \mathbf{a}_\Theta + \frac{1}{r \sin \Theta} \frac{\partial}{\partial \phi} \mathbf{a}_\phi$$

*Hint:* In both Parts (b) and (c) “additional” terms may turn out to be infinitesimal.

**4.21** The cavity of Fig. 4.5 is driven by a thin antenna extending from the bottom surface  $y = 0$ , to the top  $y = b$ , and located at  $x = a/2, z = d/12$ . The current density in the antenna is thus given by the real part of

$$\mathbf{J} = I_0 \delta(x - a/2) \delta(z - d/12) e^{-i\omega t} \mathbf{j}$$

where  $\omega$  is the driving frequency, and  $\mathbf{j}$  the unit vector in the  $y$ -direction. The charge density  $\varrho$  equals 0.

a) Show that the amplitude of the vector potential associated with the TE<sub>103</sub> mode is given by

$$\begin{aligned} C_{\text{TE}_{103}} &= \frac{\mu_0}{k^2 - k_0^2} \left( \frac{4k^2}{V k_1^2} \right) I_0 b \frac{k_1}{k} \sin \frac{k_1 a}{2} \sin \frac{k_3 d}{12} \\ &= \frac{2\sqrt{2}\mu_0 I_0}{ad(k^2 - k_0^2)} \sqrt{1 + 9a^2/d^2} \end{aligned}$$

where  $k$  is  $k_{103}$  and  $k_0$  the ratio of the driving frequency  $\omega$  to the speed of light.

- b) Discuss the concept of “resonance” in terms of this result for the amplitude, and say something about the effect of losses in the cavity walls on the result.
- c) Calculate the electric field associated with the  $\text{TE}_{103}$  mode as measured at the location of the antenna, and write an expression for the external voltage drop across the antenna.
- d) Use (c) to find the complex impedance  $Z$  for the mode. Show that for this mode that cavity’s equivalent circuit is an  $L$  and  $C$  in parallel (see Problem 2.11.). Find  $L$  and  $C$ .

The coefficients  $C_s$  are defined by the expression

$$\mathbf{A}(x, y, z, t) = \sum_s C_s \mathbf{u}_s(x, y, z) e^{-i\omega t}$$

when the  $\mathbf{u}_s$  are the modes given in (4.26) for the electric field.

- 4.22** Verify that the time average of the power radiated by a dipole antenna is given by

$$\bar{P} = \rho_0^2 \omega^4 / (12\pi \epsilon_0 c^3). \quad (4.45)$$

In spherical coordinates

$$\begin{aligned} \nabla \times \mathbf{F}(r, \theta, \phi) &= \frac{\mathbf{a}_r}{r \sin \theta} \left[ \frac{\partial}{\partial \theta} (\sin \theta F_\phi) - \frac{\partial F_\theta}{\partial \phi} \right] \\ &\quad + \frac{\mathbf{a}_\theta}{r} \left[ \frac{1}{\sin \theta} \frac{\partial F_r}{\partial \phi} - \frac{\partial}{\partial r} (r F_\phi) \right] + \frac{\mathbf{a}_\phi}{r} \left[ \frac{\partial}{\partial r} (r F_\theta) - \frac{\partial F_r}{\partial \theta} \right]. \end{aligned}$$

Show first that in the “radiation zone”,  $d \ll \lambda \ll r$  (in contrast to the “near zone”,  $d \ll r \ll \lambda$ ) the electric and magnetic fields, separately, diminish with distance from the antenna as  $1/r$ .

- 4.23** Show that Coulomb’s Law is not only valid in electrostatics, but gives the longitudinal electric field even in electrodynamics.

#### Laboratory Excercise.

Excite an EM cavity with a klystron and look for resonances. Examine the  $\mathbf{E}$  and  $\mathbf{B}$  field patterns of the resonating modes.

## 4.A Appendix

### Verification of Meaning of Vector Operators

*Gradient:* We need to show that if  $\mathbf{c}$  is a unit vector in a given fixed direction,  $\mathbf{c} \cdot \nabla \psi(x, y, z)$  gives the rate of change of  $\psi$  in the direction of  $\mathbf{c}$ .

Proceeding in two dimensions we note that the change in  $\psi$ , in going from Point 1 to Point 3 in Fig. 4.A.1 is:

$$\begin{aligned}\psi(3) - \psi(1) &= [\psi(2) - \psi(1)] + [\psi(3) - \psi(2)] \\ &= \frac{\partial \psi}{\partial x} \Delta x + \frac{\partial \psi}{\partial y} \Delta y.\end{aligned}$$

The rate of change of  $\psi$  in the direction of  $\mathbf{c}$  is

$$\frac{\psi(3) - \psi(1)}{\sqrt{(\Delta x)^2 + (\Delta y)^2}} = \frac{\partial \psi}{\partial x} \cos \theta + \frac{\partial \psi}{\partial y} \sin \theta. \quad (4.A.1)$$

On the other hand

$$\begin{aligned}\mathbf{c} \cdot \nabla \psi &= c_x \frac{\partial \psi}{\partial x} + c_y \frac{\partial \psi}{\partial y} \\ &= \frac{\partial \psi}{\partial x} \cos \theta + \frac{\partial \psi}{\partial y} \sin \theta,\end{aligned}$$

because  $\mathbf{c}$  is a unit vector.

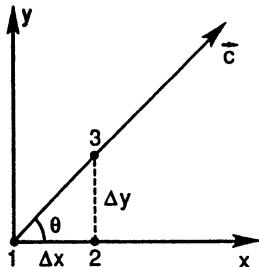
*Divergence:* We need to show that the net outward flux of the vector field  $\mathbf{V}$  from a given vanishingly small volume, per unit volume, is given by  $\nabla \cdot \mathbf{V}(x, y, z)$ .

The flux emerging from the face at  $x = \Delta x$  as shown in Fig. 4.A.2 is:

$$\left[ V_x(x, y, z) + \frac{\partial V_x(x, y, z)}{\partial x} \Delta x \right] \Delta y \Delta z$$

and the flux entering the back face at  $x = 0$  is:

$$V_x(x, y, z) \Delta y \Delta z,$$



**Fig. 4.A.1.**  $\mathbf{c} \cdot \nabla \psi$  gives the rate of change of  $\psi$  in the direction of  $\mathbf{c}$

giving the *net* flux

$$\frac{\partial V_x(x, y, z)}{\partial x} \Delta x \Delta y \Delta z .$$

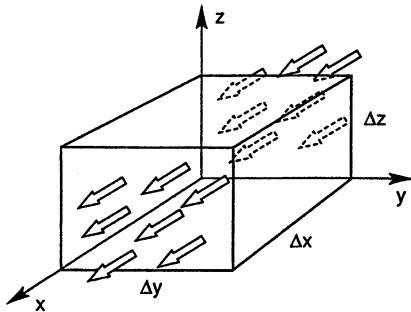
Proceeding to the net flux through the top and bottom, we find

$$\frac{\partial V_z(x, y, z)}{\partial z} \Delta x \Delta y \Delta z .$$

Finally, including the net flux through the faces at  $y = 0$ ,  $y = \Delta y$ , we find that the net outward flux per unit volume is given by:

$$\frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} + \frac{\partial V_z}{\partial z} , \quad (4.A.2)$$

as we wish to show.



**Fig. 4.A.2.** The outward flux of the vector field  $\mathbf{V}$  per unit volume is given by  $\nabla \cdot \mathbf{V}$ . Shown in the figure is the vector field  $V_x$ , with flux entering at  $x = 0$ , and leaving at  $x = \Delta x$ .

*Curl.* We need to show that the line integral of a vector field  $\mathbf{V}$  around a contour enclosing a vanishingly small area  $\Delta S$  is given by  $(\nabla \times \mathbf{V}) \cdot \Delta S$ .

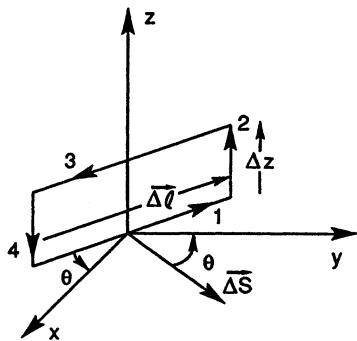
Restricting ourselves to having  $\Delta S$  in the  $x$ - $y$  plane and making an angle  $\theta$  with the  $y$ -axis, we have

$$\begin{aligned} (\nabla \times \mathbf{V}) \cdot \Delta S &= \left[ \left( \frac{\partial V_z}{\partial y} - \frac{\partial V_y}{\partial z} \right) \mathbf{i} + \left( \frac{\partial V_x}{\partial z} - \frac{\partial V_z}{\partial x} \right) \mathbf{j} \right. \\ &\quad \left. + \left( \frac{\partial V_y}{\partial x} - \frac{\partial V_x}{\partial y} \right) \mathbf{k} \right] \cdot [\mathbf{i} \sin \theta + \mathbf{j} \cos \theta] |\Delta S| \\ &= \left[ \left( \frac{\partial V_z}{\partial y} - \frac{\partial V_y}{\partial z} \right) \sin \theta + \left( \frac{\partial V_x}{\partial z} - \frac{\partial V_z}{\partial x} \right) \cos \theta \right] (\Delta z)(\Delta l) . \end{aligned} \quad (4.A.3)$$

We divide the desired line integral  $\oint \mathbf{V} \cdot d\mathbf{l}$  into four components, as in Fig. 4.A.3.

$$I_1 = \mathbf{V}(x, y, z) \cdot \Delta \mathbf{l} \equiv \mathbf{V} \cdot \Delta \mathbf{l} ,$$

$$I_2 = (V_z + \nabla V_z \cdot \Delta \mathbf{l}) \Delta z ,$$



**Fig. 4.A.3.** The line integral of a vector field  $\mathbf{V}$  around a contour of area  $\Delta S$  is given by  $(\nabla \times \mathbf{V}) \cdot \Delta S$ . The magnitude of the vector  $\Delta S$  gives the area  $\Delta l \Delta z$ , and  $\Delta S$  points perpendicular to the infinitesimal area

$$I_3 = - \left( \mathbf{V} + \frac{\partial \mathbf{V}}{\partial z} \Delta z \right) \cdot \Delta l ,$$

$$I_4 = -V_z \Delta z .$$

$$\oint \mathbf{V} \cdot d\mathbf{l} = I_1 + I_2 + I_3 + I_4 , \\ = -(\partial \mathbf{V} / \partial z) \cdot \Delta l \Delta z + (\nabla V_z \cdot \Delta l) \Delta z .$$

We note that

$$\Delta l = (-i \cos \theta + j \sin \theta)(\Delta l) .$$

Then,

$$\oint \mathbf{V} \cdot d\mathbf{l} = \left[ \left( \frac{-\partial V_x}{\partial z} i - \frac{\partial V_y}{\partial z} j \right) \cdot (-i \cos \theta + j \sin \theta) \right. \\ \left. + \left( \frac{\partial V_z}{\partial x} i + \frac{\partial V_z}{\partial y} j \right) \cdot (-\cos \theta i + \sin \theta j) \right] (\Delta z \Delta l)$$

which gives the right side of (4.A.3).

A proof with the element of area tilted with respect to the  $z$ -axis would be tedious. The symmetric form of  $\nabla$  in cartesian coordinates does make plausible that  $\nabla \times \mathbf{V}$ , a vector, is invariant under rotation or translation upon change of coordinates (the same form yields the same value at every point.  $(\nabla \times \mathbf{V}) \cdot \Delta S$ , a scalar remains equal to  $\oint \mathbf{V} \cdot d\mathbf{l}$ ).

**Problem 4.A.1.** In the analysis of a vacuum diode, Example 4.3, we obtain for the static potential the result:

$$\phi(x) = (3/2)^{4/3} (CJ)^{2/3} x^{4/3} .$$

Here, calculate  $E(x)$ ,  $\rho(x)$ ,  $v(x)$ , and check that

$$J = -\rho(x)v(x)$$

does not depend on  $x$ .

Make a plot of  $\phi(x)$ ,  $E(x)$ ,  $\varrho(x)$  versus  $x$ , and compare with similar plots when  $\varrho(x)$  vanishes, but  $\phi(0)$  and  $\phi(d)$  remain the same. What do we know about the area  $-\int_0^d E(x)dx$  in both cases?

Now, place a rectangular box extending from some point  $x$  to  $x + \Delta x$  somewhere in the range of the above graphs, i.e. for  $d > x > 0$ , with the cross-sectional area of the box parallel to the  $y - z$  plane, i.e. parallel to the planes of the cathode and anode. Draw in the electric field lines as is being done for  $\nabla_x(x, y, z)$  in Fig. 4.A.2, but inside the box as well. What is the net electric flux out of the box? To first order in  $\Delta x$ ? What is the total electric charge inside the box? Show the electric charges as they occur in the box. Do your electric field lines accurately portray the situation of the electric field changing with  $x$ ? Now, find  $E(x)$  and  $\varrho(x)$  as  $x \rightarrow 0$ , but not at  $x = 0$ ! Explain. Does the integral form of Maxwell's equation help resolve the situation at  $x = \Delta x$ ,  $\Delta x \rightarrow 0$ ?

**Problem 4.A.2.** Fluids flow in the  $y$  direction in a rectangular channel with velocities  $\mathbf{v}(x, y)$  which do not depend on the depth, i.e. on  $z$ , but possibly do on the lateral position  $x$ . Consider two situations:

- i)  $\mathbf{v}(x, y) = v(x)\mathbf{j}$ ,
- ii)  $\mathbf{v}(x, y) = v(y)\mathbf{j}$ .

Draw in the velocity field lines for both cases. Compare and contrast the two upon calculating  $\nabla \cdot \mathbf{v}$  and  $\nabla \times \mathbf{v}$ . In which case is the channel changing direction? Could one of the cases apply to a compressible fluid under acceleration? Discuss. Which flow can be described as laminar?

**Problem 4.A.3.** Show that  $\nabla \times \mathbf{V}$  is a vector.

*Hint:*  $\mathbf{i} \times \mathbf{V}(x + \delta x, y, z)$  is a vector.

## Further Reading

D.J. Griffiths: *Introduction to Electrodynamics* (Prentice Hall, Upper Saddle River, NJ, 3d. ed., 1999)

R.P. Feynman, R.B. Leighton, M. Sands: *Lectures on Physics II* (Addison-Wesley, Reading, MA, 1962)

W.K.H. Panofsky, M. Phillips: *Classical Electricity and Magnetism* (Addison-Wesley, Reading, MA, 1962)

## 5. Light – Physical Optics, Refraction

**Abstract.** The previous chapter on EM waves forms a basis for developing physical optics, which itself traditionally serves as the background to wave mechanics, (Chap. 6). In addition, in this chapter we analyze the propagation of EM waves in dielectrics, with the goal of understanding refraction and reflection at interfaces, the central phenomenon in geometric optics. The chapter begins with a discussion of the generation of light as the outcome of quantum transitions within matter and the role of the Heisenberg uncertainty principle in limiting coherence is pointed out. Analysis of diffraction by a single slit using Fresnel zones, and the consequences for the resolution of illuminated objects follows. Special considerations applying to X-ray diffraction by crystals (three dimensional gratings) completes the discussion of physical optics. Propagation of plane EM waves in insulating dielectrics is treated. Lastly, extension of the Green's function method of solving inhomogeneous wave equations to three dimensions provides a rigorous basis for Fresnel analysis of diffraction. Worked examples include interference of light from two slits, as well as details of reflection and refraction at an interface. There are problems dealing with the Einstein photoelectric effect, the Huygens principle, the square aperture, gratings, Rayleigh's criterion for resolution, Debye-Scherrer X-ray diffraction, and attenuation in metals.

### 5.1 The Nature and Generation of Light

The human eye is sensitive to electromagnetic radiation with wavelengths in the range of 6500 Å in the red to 4500 Å in the violet. ( $1 \text{ \AA} = 10^{-10} \text{ meters}$ ). Certain phenomena characteristic of wave motion in general, and EM waves in particular, namely *diffraction* and *interference*, have been studied intensively in the visible region of light. These studies form the subject matter of physical optics, and much of this chapter will be devoted to them.

According to classical physics, that is, physics as it stood prior to the development of quantum mechanics in the twentieth century, all EM radiation is generated by accelerating electric charges. For light in the visible region the charges are, generally, electrons attached to atoms. In the same context one can think of  $\gamma$ -rays as generated by the charges of accelerating nuclear

particles comprising the nuclei, infrared rays by accelerating ions, and, as already mentioned, radio waves by the macroscopic charge densities forming in antennae. In every case we can think of the motion of these charges as oscillations about an equilibrium point, giving us the so-called electromagnetic oscillators. The frequency of the EM radiation then corresponds to the frequency of the charge oscillation.

In modern physics the situation is described somewhat differently. Nevertheless, despite the seemingly disparate nature of the classical and modern concepts, the modern merge into the classical under certain conditions in accord with the “correspondence principle”. This principle was first enunciated by Niels Bohr; be learning more about his many contributions. In quantum physics particles occupy certain *states*; if *stationary*, these states have a definite and constant energy associated with them. If a particle makes a transition from a state whose characteristic energy is  $E_a$  to one whose energy is  $E_b$ , the energy difference ( $E_a - E_b$ ) is, assuming it to be positive, emitted in some form, quite often as EM radiation. What is the frequency of this radiation?

In 1905 Albert Einstein gave his famous interpretation of the photoelectric effect. He concluded that EM radiation is given off as discrete entities, quanta, and that each quantum of radiation, called a photon, will have associated with it an energy  $h\nu$ , where  $\nu$  is the frequency of the radiation in Hertz, and  $h (= 6.624 \times 10^{-34} \text{ joule s})$  is called Planck's constant, after Max Planck. Planck had been able a few years earlier, in 1900, to fit the so-called black-body radiation with a formula that required one new hypothetical physical constant,  $h$ , which he was able to roughly evaluate from the experimental data. Einstein's hypothesis gave a clear meaning to Planck's findings.

It turns out rather simply that when a particle makes a downward transition from an energy level  $E_a$  to a level  $E_b$  that it generally emits only a single photon, one that has the energy relationship

$$E_a - E_b = h\nu_0 , \quad (5.1)$$

giving us at once a frequency,  $\nu_0$  for the emitted radiation.

In an isolated atom the characteristic energies of the various states, the various electronic *energy levels*, are well separated. It follows that the emitted radiation occurs at only certain discrete frequencies. However, when the atoms become associated into a crystal, for example, a crystal of tungsten, the discrete levels widen into continuous energy *bands*. We then get a continuous *spectrum* with all frequencies, rather than the discrete spectra characteristic of gases, i.e., of isolated atoms.

Thus, on the one hand light comes in discrete “particles” called photons. Yet, clearly on the other, it also has a wave nature, how otherwise could we speak of a frequency  $\nu_0$ ? Our study of the Fourier representation in Chap. 1 will help us to resolve this paradox in large measure. Further clarification will have to await the ensuing chapter, where we broaden our discussion to include matter waves.

We naturally suppose that the radiation is emitted at some definite time,  $t = 0$ . However, it is more fruitful to suppose that a pulse of radiation is emitted which occupies some finite time interval about  $t = 0$ , let us say, from  $-\Delta t/2$  to  $+\Delta t/2$ . The magnetic field, for example, associated with the pulse is then present during this time interval, as shown in Fig. 5.1. Let us represent this field as

$$B(t) = \cos \omega_0 t \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} B_{\omega} e^{i\omega t} d\omega \quad (5.2a)$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} B_{\omega - \omega_0} \cos(\omega t) d\omega \quad (5.2b)$$

$$\omega_0 = 2\pi\nu_0 , \quad (5.2c)$$

$\nu_0$  being given in (5.1). In (5.2a) we have separated out the factor giving oscillations with the frequency  $\nu_0$ . The remaining factor,

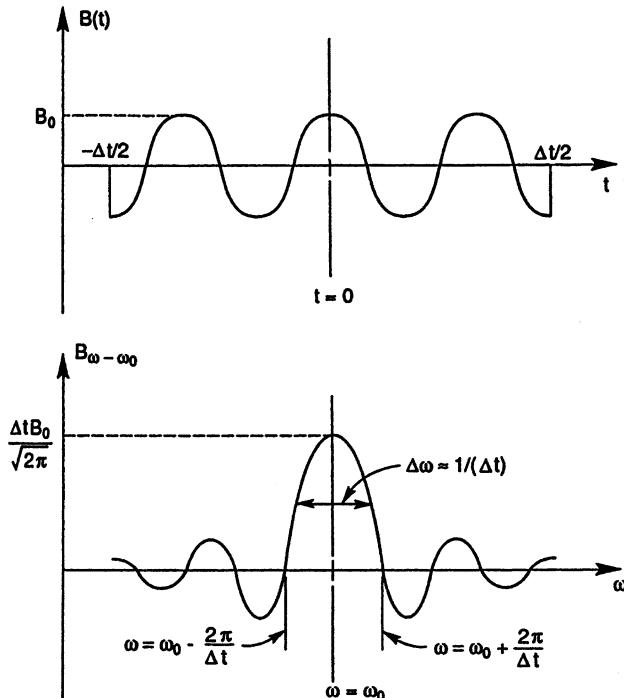


Fig. 5.1. A light signal with the magnetic field  $B(t)$  occupying a finite time  $\Delta t$ .  $B_{\omega - \omega_0}$  is the cosine Fourier transform of its square wave envelope

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} B_\omega e^{i\omega t} d\omega$$

represents the square-wave envelope extending from  $-\Delta t/2$  to  $\Delta t/2$  with amplitude  $B_0$ . We already know from Example 1.1 that the cosine Fourier transform of  $B(t)$ ,  $B_\omega - \omega_0$ , will satisfy

$$B_{\omega-\omega_0} = B_0 \sqrt{2/\pi} \sin [(\omega - \omega_0)\Delta t/2] / (\omega - \omega_0) , \quad (5.3)$$

(Fig. 5.1). A more realistic representation of the radiated signal would be based on the statement that the probability of finding the particle in its upper excited state goes like  $e^{-\gamma t}$ , where  $1/\gamma$  is the “lifetime” of the excitation. This would result in a signal with the same time dependence as that of damped oscillations (Problem 5.3).

Actually, we are not interested here in the detailed structure of either  $B(t)$  or its transform  $B_{\omega-\omega_0}$ , but rather in the general feature universally true of the Fourier relation: If  $B(t)$  occupies a region in time of roughly  $\Delta t$ ,  $B_{\omega-\omega_0}$  will occupy a region in angular frequency whose width  $\Delta\omega$ , about  $\omega_0$  is related to  $\Delta t$  by the approximate relation

$$\Delta\omega\Delta t \simeq 1 . \quad (5.4)$$

We have already indicated this relationship in Chap. 1, see, in particular, Example 1.1: a smooth broad function has a narrow Fourier transform, and vice versa.

To recapitulate, light is generated by the downward transition of an electron from one energy level to a lower one. (Absorption accompanied by an upward transition is also possible. In quantum mechanics the peak of the exponential Fourier transform of  $B(t)$  at  $-\omega_0$  then comes into play.) The resulting pulse of light occupies some finite time interval, which is related reciprocally to the range of harmonic frequencies needed to represent the pulse in time. But how does (5.1), which links the emitted frequency to energy levels, fit in?

We shall learn in the next chapter that only perfectly stable systems, systems that do not change at all in time, are associated with perfectly sharp energy levels, i.e., are found in stationary states. This will turn out to be in accordance with the most comprehensive principle of quantum mechanics, the “Heisenberg uncertainty principle”. However, the atoms of, say, a crystal of tungsten are not perfectly stable. Not only are they disturbed by the thermal activity of the crystal, but the very radiation process itself diminishes the stability of the upper levels. Let us, therefore, assume that the width at the level of  $E_a, \Delta E$ , is actually some definite quantity, often in practice on the order of  $10^{-20}$  joules, while the width of the lower level may be neglected. Then, by (5.1) the spread in emitted frequencies is simply

$$\Delta\nu = \Delta E/h, \quad \Delta\omega = 2\pi \Delta E/h ,$$

and from (5.4) the duration of the pulse, which we can take to represent *the uncertainty in the time of its emission*, is given by

$$\Delta t \simeq \frac{1}{\Delta E} \cdot \frac{\hbar}{2\pi}$$

or

$$\Delta E \Delta t \simeq \hbar/2\pi . \quad (5.5)$$

This relation is the Heisenberg uncertainty principle as it applies to radiation. Does it have any practical consequences for the study of diffraction and interference upon which we are about to embark?

Indeed, yes. The phenomena of diffraction and interference come about as a result of a recombination of different portions of a light beam. We naturally write the magnetic (or electric) field for such a beam as

$$B(\omega) = B_0 \sin(kx - \omega t + \delta_\omega) .$$

Here  $\delta_\omega$  is the phase, which we are showing explicitly. Some reflection will show that  $\delta_\omega$  is a determined quantity as long as we are concerned with light originating from one particular radiation process, but we do not expect or get any relation between this phase and the phase pertaining to light at a given frequency that originated from a different transition. Thus, although we will be observing the accumulation of effects from many radiation processes, these will be adding together in a random fashion. The special relationships we will be studying will apply only to radiation originating at most from a *coherence length*

$$l \simeq c\Delta t \simeq \frac{\hbar c}{\Delta E} , \quad (\hbar = \hbar/2\pi) , \quad (5.6)$$

i.e., from one radiation process.

In view of its feature of random phases we refer to ordinary light as *incoherent* radiation. The reader is probably familiar with the fact that there are devices, *lasers*, that produce highly coherent radiation. Interference and diffraction become much more striking effects when one works with laser light rather than with traditional sources. The phenomena first discovered by physicists in the early 19th century, and studied by generations of students since, become really vivid only now in the quantum age.

## 5.2 Diffraction

So far our work has been confined primarily to closed systems such as lumped oscillators, finite stretched strings, closed cavity resonators. However, we have also briefly seen a different type of problem, an open system with a steady flux. Examples are the wave traveling down an infinite string in Problem 3.14, or the traveling plane EM waves of the previous chapter. These problems generally have a common simplifying feature; they are steady state problems.

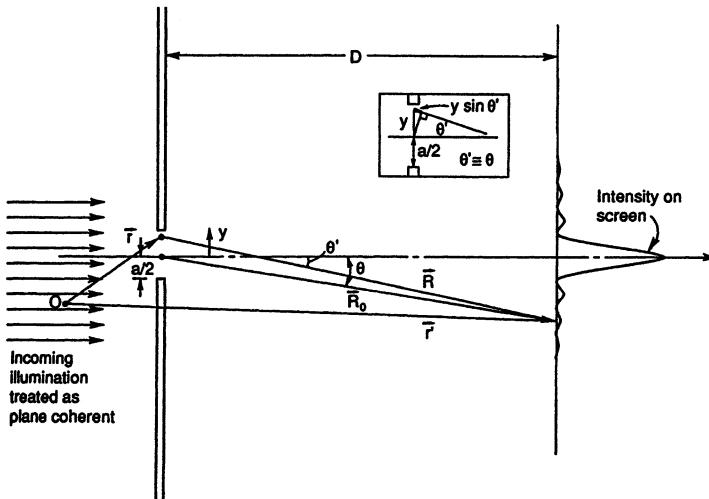
The flux or flow keeps coming steadily, it never starts or stops. The transients associated with the startup are assumed to be long past. It would seem that there is not much to be learned about such situations, which is true until the “beam” is interfered with in some way, some obstacle such as a narrow slit or a reactive body is placed in its path. We then get a so-called “scattering problem”. Many chemical reactions can be viewed as scattering problems, and so can the problems of diffraction and interference of physical optics.

The fields comprising the flux of energy obey wave equations. The method of solution often resorted to is the Green’s function method, already introduced in previous chapters. The Green’s function is the solution of the wave equation that corresponds to a “point” source in time or space or both, appropriate boundary conditions also being fulfilled. Once the Green’s function has been found the problem can, at least in principle, be solved easily, with the magic rule already illustrated in Chap. 3, where we only needed to work in one dimension. In Sect. 5.5 the reader will find a complete analysis illustrating the Green’s function method in physical optics.

Here we will content ourselves with working with not much more than a rather old principle, formulated by C. Huygens around 1670. The mathematically complete formulation from which the principle can be derived, though, is just the Green’s function method of theoretical physics. Huygens’ principle may be stated as: each point on a wave front acts as a point source of a secondary disturbance, of a spherical wavelet, also traveling with velocity  $c$ . The field at a given point some time later is then the sum of the fields of each of these secondary disturbances, their envelope being the new wave front.

Consider now the problem of a plane coherent beam of light passed through a rectangular aperture, Fig. 5.2. How will a screen placed a distance  $D$  from the aperture be illuminated? If we were dealing with a beam of fast particles moving according to Newtonian mechanics, then the illumination would be confined to a rectangular area corresponding to the aperture opening, the geometric shadow. The light, on the other hand, will spread out somewhat and, further, the intensity of the illumination will vary rapidly on the screen according to a complicated function. Such behavior, which includes the tendency to bend around corners, characterizes all wave phenomena, and is called *diffraction*. As we will see, one can determine the wavelength associated with the beam from the pattern of illumination on the screen, i.e., the *diffraction pattern*.

We assume that the slit is very long in the direction out of the plane of Fig. 5.2 (horizontal), so that, in fact, we will be dealing with a one dimensional problem that of determining the variation in the amplitude of the fields as a function of the (vertical) displacement on the screen. In view of the Huygens principle we shall divide the aperture into elements of area  $(dxdy)$ ,  $x$  being the coordinate which measures along the extended length of the aperture (i.e., out of the plane of Fig. 5.2),  $y$  the vertical coordinate. We consider each element to be the source of a secondary wave of light.



**Fig. 5.2.** Single slit diffraction.  $\mathbf{R} = \mathbf{r}_1 - \mathbf{r}$  as measured from arbitrary origin O.  $|\mathbf{R}| \simeq |\mathbf{R}_0| + y \sin \theta$ . Slit width greatly exaggerated

How can we describe the electric (or magnetic) field produced by the secondary elemental sources analytically? According to Huygens' original principle the secondary wavelets are taken as isotropic radiators in the forward direction. This seems rather reasonable on an intuitive basis, since, first and foremost, the principle must be capable of explaining the progression of an ordinary plane wave. (The complete Green's function analysis yields an obliquity factor that disallows backward radiation. In view of the small angles involved the factor can be taken as constant in the forward direction, see Sect. 5.5.) In addition, in any optical system the polarization of the incoming radiation will usually remain largely unaffected by the system itself, with the result that the polarization of the transmitted radiation need not be particularly specified. Thus, as a good approximation, one ends up representing the electric or magnetic fields by a scalar field, the so-called *optical disturbance*, which is taken as radiating equally in all directions from the secondary sources. At the same time one bears in mind the considerations of the previous section, that light interferes with itself in discernible fashion only within a coherence length.

As a result, we restrict ourselves only to the magnitude of the electric field associated with the spherical wavelet radiated from the element at  $\mathbf{r}$ , and write for this optical disturbance measured at the point  $\mathbf{r}'$  on the screen the proportionality

$$dE \propto \frac{e^{ikR}}{R} dx dy \quad (5.7a)$$

$$\mathbf{R} = \mathbf{r}' - \mathbf{r} . \quad (5.7b)$$

The constant of proportionality (which in any case can not be obtained directly, see the appendix), does not play an essential role in this analysis.

The expression (5.7a) represents a spherical wave of wave vector  $k$ . We notice that the points of constant phase lie on concentric spherical surfaces with the point  $\mathbf{r}$  at the center, and that the radiation flux per unit area, which is proportional to the square of the amplitude, satisfies an inverse square law: the same power is crossing every spherical surface, as required by energy conservation.

To find the total amplitude of the field at the point  $\mathbf{r}'$  on the screen we add the contributions from all the elements of area comprising the aperture. We can simplify further by accepting the fact that in optical experiments the variations  $\delta R$  in  $R$  in the denominator of (5.7a) are negligible compared to  $R$ , while variations in the phase factor  $kR$  equal to  $2\pi\delta R/\lambda$ , are rapid;  $\delta R$  is not small compared to the wavelength  $\lambda$  of the light. Accordingly we write

$$E(\mathbf{r}') \propto \int \int e^{ikR} dx dy, \quad \mathbf{R} = \mathbf{r}' - \mathbf{r}(x, y). \quad (5.8)$$

We have already said that the aperture is long in the horizontal ( $x$ ) direction and that we can treat the problem as a one dimensional one. (If we are at the horizontal middle of the pattern on the screen proceeding vertically we can always add contributions (phasors) from two segments placed symmetrically about  $x = 0$  on the aperture. This will yield the same net phase as that due to the element of the aperture at  $x = 0$  for the given value of  $y$ . As we depart from the center of the screen the procedure breaks down since not all the elements can now be paired. However, the assumption of a long slit, viz., the tacit omission of end effects, yields a constant intensity in the  $x$ -direction. The interested reader is referred to Problem 5.7, the square aperture.)

If we specify the position of the chosen point  $\mathbf{r}'$  on the screen by the angle  $\theta$  (Fig. 5.2) then we may write to good approximation

$$R = R_0 + y \sin \theta. \quad (5.9a)$$

[This approximation becomes exact if  $a/D \rightarrow 0$ , see Fig. 5.2, or if we place a lens next to the aperture which would exactly focus the beam specified by  $\theta$  at the point  $\mathbf{r}'$ . In practice one either does use a lens (Fraunhofer diffraction), or one chooses  $a/D$  small (Fresnel diffraction).]

Combining (5.8) and (5.9) we finally have

$$B(\mathbf{r}') \propto e^{ikR_0} \int_{-a/2}^{+a/2} e^{iky \sin \theta} dy.$$

Integrating, one finds

$$B(\mathbf{r}') \propto e^{ikR_0} \left| \frac{2 \sin(ka \sin \theta / 2)}{k \sin \theta} \right|. \quad (5.9b)$$

It is customary to let

$$\alpha = \frac{ka \sin \theta}{2} = \frac{2\pi}{\lambda} \frac{a \sin \theta}{2}, \quad (5.9c)$$

giving

$$B(r') \propto e^{ikR_0} \left| \frac{\sin \alpha}{\alpha} \right|. \quad (5.9d)$$

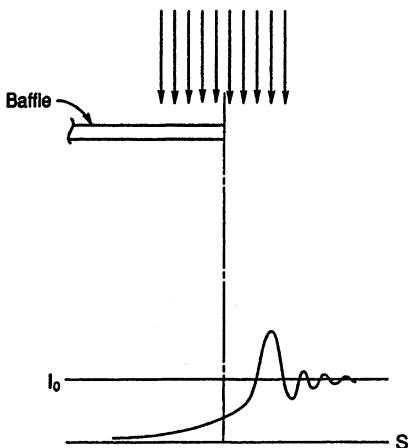
The intensity of the illumination,  $I$ , power/unit area (Problem 4.8) is proportional to the square of the magnetic, or electric, field. The intensity will be most intense at the center of the pattern, i.e., at  $\alpha = 0$ . We note in passing that  $\alpha$ , as given by (5.9c), represents the optical phase difference between the ray coming from the upper end of the aperture and the one coming from its center, Fig. 5.2. When  $\alpha = 0$ , the factor  $\sin \alpha / \alpha$  and its square have their maximum values, namely unity. Squaring  $B(r')$ , we may write for the intensity at  $r'$ , that is, at  $\theta$ ,

$$I(\theta) = I_0 \left( \frac{\sin \alpha}{\alpha} \right)^2 \quad (5.10)$$

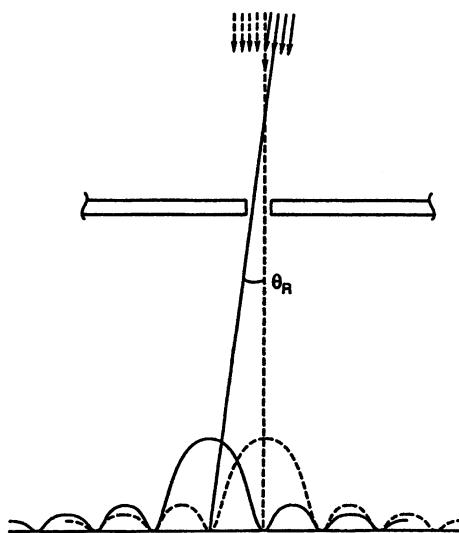
where  $I_0$  is the maximum intensity at the center of the pattern.  $I(\theta)$ , the one slit diffraction pattern, is shown in Fig. 5.2.

The diffraction patterns of many geometries, besides that of a single long slit, have been thoroughly studied. In the problems we look at the pattern of the finite rectangular slit. Very well known is the half plane pattern shown in Fig. 5.3. Here we see particularly graphically that light will bend around corners.

What if we have more than one opening? The interference from two slits is discussed in Example 5.1. The illuminating patterns become yet more complicated (see Figs. 5.9 and 5.11), and are referred to as *interference patterns*.



**Fig. 5.3.** Diffraction by a half plane. Light bends around corners.  $I_0$  is the light intensity on the screen  $S$  far from the shadow



**Fig. 5.4.** Diffraction introduces a limitation on information that can be imparted by signals of a wave nature

One can also have three or more slits, or many thousands. In the latter case we talk about a *diffraction grating*, a scientific instrument which the reader will find analyzed in Problems 5.9 and 5.10.

Diffraction brings about a fundamental limitation on the information that can be imparted by signals of a wave nature. In Fig. 5.4 we show the diffraction pattern produced when light from two sources placed close together is passed through a rectangular slit. We see that as the two sources approach each other the two patterns will merge, and it shall no longer be clear that there is anything other than one extended source of illumination. If we define  $\theta_R$  as the smallest angle separating the two sources which allows us to conclude that the sources are distinct, then we have approximately that

$$\theta_R \simeq \lambda/a , \quad (5.11)$$

since the first minimum of the diffraction pattern occurs at

$$\alpha = \frac{2\pi}{\lambda} \frac{a \sin \theta}{2} = \pi, \quad \sin \theta = \frac{\lambda}{a} .$$

In (5.11) we have made use of the so-called *Rayleigh criterion* (Baron Rayleigh 1842–1919) which suggests that two objects may be seen as distinct as long as the maximum of one pattern lies clear of the first minimum of the other.

When viewing objects we are generally restricted to light received through an aperture, for example, the lens of the human eye. When these apertures are circular one needs to carry out the corresponding analysis for this type of geometry. This involves functions that may not be familiar to the reader (Bessel functions) and we shall not carry this analysis through here. Let

us merely note that the pattern of a circular aperture is a series of concentric rings, with the intensity along any radius having a variation rather resembling the lateral behavior we have found with the long rectangular slit. The Rayleigh criterion, which actually was originally formulated for circular apertures, becomes

$$\theta_R = 1.22\lambda/d,$$

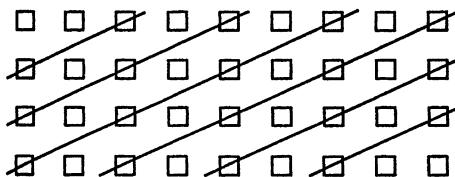
where  $d$  is the diameter.

### 5.3 X-ray Diffraction

The region of 10 to 1.10 Å of the electromagnetic spectrum, i.e., about 1/1000 of the wavelength of visible light, is associated with X-rays. While sharp X-ray lines are obtained when electrons located deep inside atoms of the heavier elements undergo transitions between their energy levels, the *continuous* X-ray spectrum, which will be of interest here, is emitted by high speed electrons ( $5 \times 10^4$  eV) when they are decelerated by their interaction with the nuclei of a target material. The radiation is sometimes called *Bremsstrahlung*, “bremsen” meaning “to brake” in German and “strahlung”, radiation.

We have already mentioned *interference* by very many slits, brought about by so-called *diffraction gratings*. In 1912, Max von Laue thought that the regular array of atoms in a crystal, being separated by distances on the order of the magnitude of X-ray wavelengths, i.e., could be made to act as a three dimensional diffraction grating. Subsequent experiments by Friederich and Knipping confirmed this hypothesis.

When electromagnetic radiation impinges on an atom the electrons of the atom will either undergo transitions to other energy levels, as discussed in Sect. 5.1, or they can scatter the incoming light (re-emit the light in a different direction) without changing their energy levels at all. It is the latter, *elastic* scattering, that gives rise to X-ray diffraction. Let us see how the latter process can be understood.



**Fig. 5.5.** Cross-sectional view of equivalent planes for a cubic structure. The squares mark the unit cells, the diagonal lines are one of many possible choices for the equivalent planes

In Fig. 5.5 we show a simplified representation of a crystal as comprised of a regular array of small clusters of atoms, the so-called *unit cells*. In many crystals the unit cell has two or more atoms, but in the simplest crystals, for example, a crystal of metallic sodium, there would be only one atom. In any

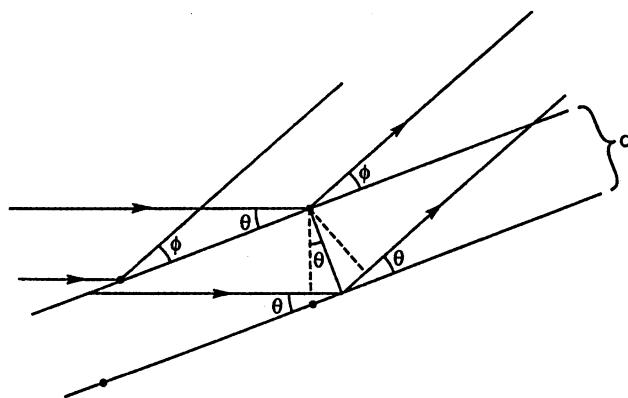
case the diffraction is brought about by the collective response of the unit cells, each viewed as a single entity, rather than by the individual atoms. Thus, in Fig. 5.6 the black dots represent unit cells. It is possible to place all the unit cells on a series of equivalent planes, as suggested in Fig. 5.5. The planes which are possible will depend on the crystal structure. Now the unit cells belonging to a single plane can act cooperatively in scattering the incoming X-ray beam. We see that if we focus our attention on that portion of the scattered radiation for which the angle of incidence is equal to the angle of reflection,  $\theta = \phi$  in Fig. 5.6 (in X-ray diffraction one works with the angle complementary to that chosen in optics), then the radiation from all atoms in the plane will be in phase. (Successive rays might also differ by one or more whole wavelengths. This is one of several possible complicating features on the conditions for reinforcement that we ignore here.) Accordingly, one should think of a given plane as a single reflecting mirror for the incoming X-rays.

Now, how do the light components specularly reflected from successive crystal planes interfere with each other? We see from Fig. 5.6 that these beams can likewise have phase coherence providing only that

$$2d \sin \theta = m\lambda, \quad m = 1, 2, 3. \quad (5.12)$$

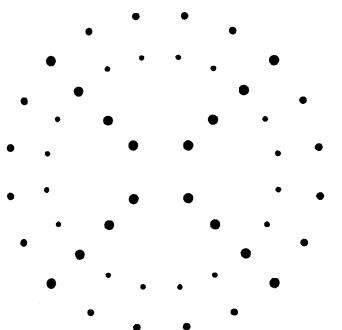
Equation (5.12) is the famous Bragg's law, named after Sir William Lawrence Bragg. It says that the difference in the path length in a crystal,  $2d \sin \theta$ , must be an integral number of wavelengths long for observable coherence.

In an experiment, a narrow beam of X-rays with a continuous spectrum is directed at a firmly mounted single crystal. The source and detector are sufficiently distant from the crystal to allow *plane wave* representation. For a given set of planes  $\theta$  is fixed. There will then in general be, for a given  $m$ ,



**Fig. 5.6.** Although atoms scatter X-rays in all directions, when  $\theta = \phi$ , X-rays from successive atoms in an equivalent plane are all in phase. Similarly, when  $2d \sin \theta = m\lambda$ , X-rays from successive planes are likewise in phase

some wavelength in the narrow incoming X-ray beam for which (5.12) will be satisfied. The narrow diffracted X-rays show up as a Laue spot on a photographic film. Usually, there will be one spot for each set of planes, giving rise to a regular pattern, Fig. 5.7. Inspection of the pattern at once gives direct information about the symmetry of the crystal target and, incidentally, allows one to locate the direction of the crystal axis. It is also possible to do the experiment with monochromatic X-ray beams and often more readily available powdered rather than single crystal targets. This is the so-called Debye-Scherrer method (Problem 5.12). We see that X-ray diffraction offers an unequivocal way of verifying the crystallinity and determining the structure of a given material.



**Fig. 5.7.** Schematic Laue pattern for an aligned crystal

## 5.4 EM Waves in Dielectrics. Refraction

The reader will be familiar with the concept of electric *polarization of materials*, that is, the separation of the positive and electric charges comprising any material which occurs when the material is exposed to an electric field. The quantity characterizing this phenomenon is the polarization vector field  $\mathbf{P}(x, y, z, t)$ , which is defined as the dipole moment per unit volume at the point and moment of time in question. The central result concerning  $\mathbf{P}$  is that

$$\nabla \cdot \mathbf{P} = -\varrho_{\text{pol}} \quad (5.13)$$

where  $\varrho_{\text{pol}}$  is the charge density arising from the polarization (Problem 5.14).

Restricting ourselves as usual to linear phenomena we write, breaking up all fields into Fourier components in space and time,

$$\mathbf{P}(\mathbf{k}, \omega) = \varepsilon_0 \chi(\mathbf{k}, \omega) \mathbf{E}(\mathbf{k}, \omega) \quad (5.14)$$

where  $\chi$  is called the electric susceptibility. Since  $\chi$  is a real (see below) scalar, we are assuming that  $\mathbf{P}$  and  $\mathbf{E}$  are parallel, as they usually tend to be, and

in phase. Physically, we differentiate between two different origins of the polarization: the first is the motion of electrons relative to the fixed centers of the ions of the material (electronic polarization); the second involves the motion of the ion cores themselves (ionic and dipolar polarization, the latter requiring the existence of internal dipolar molecules). The frequency response of these different mechanisms will be different, the electronic polarization will be active up to a higher cutoff frequency than for the other type of mechanism.

The simple model of polarization in Problem 2.10 based on an ensemble of forced harmonic oscillators illustrates in some detail the origin of the frequency dependencies. Thus, it comes about that  $\chi(\mathbf{k}, \omega)$  is indeed a function of the frequency,  $\omega$ . The wave vector,  $\mathbf{k}$  dependence is actually only of interest in metals, where one can speak of the conduction electrons as forming an independent electron gas. In this section we shall consider only insulators, and omit further mention of the wave vector dependence of  $\chi$ .

The reader will also be familiar with the *displacement field*  $\mathbf{D}(\mathbf{r}, t)$ , very useful in all problems involving *dielectrics* i.e., materials that polarize. The definition of  $\mathbf{D}(\mathbf{r}, t)$  is

$$\mathbf{D}(\mathbf{r}, t) \equiv \epsilon_0 \mathbf{E}(\mathbf{r}, t) + \mathbf{P}(\mathbf{r}, t) . \quad (5.15)$$

Again, resorting to Fourier components, but omitting any wave vector dependence, we have

$$\mathbf{D}(\omega) \equiv \epsilon(\omega) \mathbf{E}(\omega) . \quad (5.16a)$$

Lastly, from (5.14),

$$\epsilon(\omega) = \epsilon_0 [\chi(\omega) + 1] , \quad (5.16b)$$

where  $\epsilon(\omega)$  is the *dielectric constant* of the material.

The central problem we wish to answer in this section is how do electromagnetic waves propagate in insulators; how does the propagation differ from that in empty space described in Sect. 4.3? If we have  $N$  dipole moments  $p_x$ , with say  $p_x = qx$ , then we shall also have a current density  $J_x = qNdx/dt$ , as well as a polarization  $P_x = Nqx$  so that

$$\mathbf{J}_{\text{pol}} = d\mathbf{P}/dt . \quad (5.17)$$

In other words, time varying polarizations produce currents, and set up magnetic fields. Suppose we are interested in the situation where the only sources, charge and current densities, are due to the polarization. Then, from (5.17) the Maxwell equation (4.12) becomes in the dielectric

$$\nabla \times \mathbf{B} = \mu_0 \left[ \frac{\partial \mathbf{P}}{\partial t} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right] . \quad (5.18)$$

By the same token (4.9) becomes

$$\nabla \cdot (\mathbf{E} + \mathbf{P}/\epsilon_0) = 0 , \quad (5.19)$$

in view of (5.13). The remaining two Maxwell equations do not change:

$$\nabla \times \mathbf{E} = -\partial \mathbf{B} / \partial t , \quad (4.10)$$

$$\nabla \cdot \mathbf{B} = 0 . \quad (4.11)$$

Following Sect. 4.3 we can easily solve these equations. Since  $\mathbf{P}$  is taken parallel to  $\mathbf{E}$  and has the same wave vector dependence we can again expect from (5.19) transverse solutions, i.e.,

$$\nabla \cdot \mathbf{P} = \nabla \cdot \mathbf{E} = 0 . \quad (5.20)$$

If we take the curl of (4.10) and use the vector identity

$$\nabla \times (\nabla \times \mathbf{E}) = \nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E}$$

we get, remembering (5.20)

$$-\nabla^2 \mathbf{E} = -\partial / \partial t (\nabla \times \mathbf{B}) .$$

Using (5.18) and again (5.20), we find

$$\nabla^2 \mathbf{E} - \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} - \frac{1}{\epsilon_0 c^2} \frac{\partial^2 \mathbf{P}}{\partial t^2} = 0 . \quad (5.21)$$

Lastly, if we have to contend with only one frequency  $\omega$  at a time, then we can combine (5.14) and (5.16a) to get

$$\mathbf{E}_\omega(\mathbf{r}, t) + \mathbf{P}_\omega(\mathbf{r}, t) / \epsilon_0 = [\epsilon(\omega) / \epsilon_0] \mathbf{E}_\omega(\mathbf{r}, t)$$

where the subscript reminds us that all time dependences will actually be harmonic with the given frequency. Consequently, we can write (5.21) as

$$\nabla^2 \mathbf{E}_\omega - \frac{\epsilon(\omega)}{\epsilon_0 c^2} \frac{\partial^2 \mathbf{E}_\omega}{\partial t^2} = 0 . \quad (5.22a)$$

The same steps can be repeated to get a similar result for  $\mathbf{B}$ , this time by taking the curl of (5.18):

$$\nabla^2 \mathbf{B}_\omega - \frac{\epsilon(\omega)}{\epsilon_0 c^2} \frac{\partial^2 \mathbf{B}_\omega}{\partial t^2} = 0 . \quad (5.22b)$$

Equations (5.22a,b) show us that we can expect the same propagating plane waves in a dielectric as we obtained for free space in Sect. 4.3. The main difference is that the wave speed is modified from  $c$ , the speed in free space, to

$$v(\omega) = \sqrt{\epsilon_0 / \epsilon(\omega)} c ; \quad (5.23)$$

$v(\omega)$  will always be numerically less than  $c$ .

Equation (5.23) is of the greatest practical importance. For if a beam of light traverses the interface from one material with a given value of the dielectric constant to another with a different value, then as a consequence of (5.23), the direction of the beam is changed. The light beam is said to have been *refracted*, from the latin “*refractus*” (broken off). The phenomenon of refraction underlies the functioning of lenses and prisms; hence its practical

importance. The ratio  $c/v(\omega)$  is, accordingly, given a special name, the *index of refraction*,  $n(\omega)$ :

$$n(\omega) \equiv c/v(\omega) = \sqrt{\epsilon(\omega)/\epsilon_0(\omega)} . \quad (5.24)$$

For light refracted by prisms the frequency dependence of the index of refraction leads to decomposition according to color (wavelength).

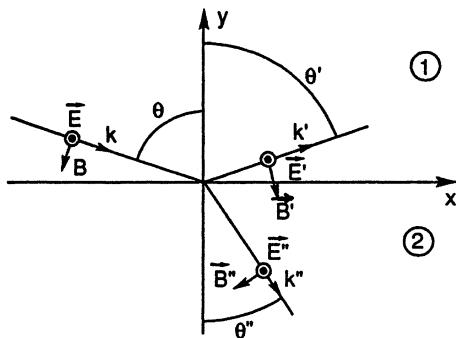
We shall analyze the problem of a beam impinging on a plane interface between two dielectrics. This will be done on the assumption that there is no loss of energy (attenuation) associated with the propagation of the light beams. In the preceding development the tacit assumption has been made that the electric susceptibility  $\chi(\mathbf{k}, \omega)$  and, hence, the index of refraction are real. However, this need not always be the case; if there is a phase lag between the applied electric field  $\mathbf{E}$  and the responding polarization (charge displacement)  $\mathbf{P}$  then power will be dissipated in the dielectric just as in the driven harmonic oscillator when damping leads to a phase difference (other than 0, or possibly  $\pi$ ) between the applied force and the displacement. One can then still use (5.14) and (5.16) to represent the situation, only now  $\chi(\omega)$  will be a complex number (Problem 2.10). Thus there arises the possibility of generalizing to a complex index of refraction. However, for the purposes of this section the electric susceptibility  $\chi(\omega)$ , the dielectric constant  $\epsilon(\omega)$ , and the index of refraction  $n(\omega)$  can be taken as real.

$\chi(\omega)$ , however, can be positive or negative. In a driven harmonic oscillator model for polarizing electric dipoles, a model which, as it turns out, has considerable physical reality, the sign of  $\chi(\omega)$  changes as  $\omega$  goes through its resonance value  $\omega_0$ . (In practice, there are a number of resonances corresponding to the different polarization mechanisms.  $\chi(\omega)$  goes negative for  $\omega$  above the highest of these.) The change of sign in  $\chi(\omega)$  has the consequence that  $n(\omega)$  becomes less than unity at frequencies above  $\omega_0$ , leading, by (5.24), to a velocity  $v(\omega)$  which exceeds the velocity of light in free space,  $c$ . It takes a more detailed consideration of the propagation of wave phenomena to see how the dictum of relativity that nothing can travel faster than  $c$  is preserved. Basically, the  $v(\omega)$  we are finding is called the phase velocity, the velocity of a single harmonic wave,  $\cos(kx - \omega t)$ , which, as we have seen in Chap. 1, implies that the wave stretches from minus to plus infinity in time and space. To convey some information we must combine many waves (invoke a *wave packet*). This *packet* will travel at its “group velocity”, which will always be less than  $c$ . An analytic exposition of wave packets and group velocity is best deferred until Chap. 7, where we consider the effects of dispersion fully, and after we have actually worked with wave packets, which we do in Chap. 6.

We now address the problem of what happens when a beam of light in a medium of refractive index  $n_1$  impinges upon a medium of index  $n_2$ ,  $n_1 \neq n_2$ , the two media separated by a plane surface, the  $x$ - $z$  plane in Fig. 5.8. As mentioned, geometric optics is based directly or indirectly on an understanding of this physical problem. Its solution depends on a knowledge of the boundary conditions at the surface. Directly from Chap. 4 we know

that in the absence of surface currents, we are dealing with nonmagnetic insulators, that  $\mathbf{B}_1, \mathbf{B}_2$ , the magnetic inductions in the two media, respectively, will be the same just above and just below the surface, as will be the tangential components of the electric field,  $E_x$  and  $E_z$ . However, because the polarization of the media leads to surface charges, the normal components of the electric field  $\mathbf{E}$  are not, in general, continuous across the boundary, and one proceeds best at this point by turning to the displacement field  $\mathbf{D}(\mathbf{r}, t)$  defined in (5.15). In the absence of charges external to those comprising the medium, (5.19) holds and we find

$$\nabla \cdot \mathbf{D} = 0. \quad (5.25)$$



**Fig. 5.8.** Incoming ( $\mathbf{k}$ ), reflected ( $\mathbf{k}'$ ), and transmitted ( $\mathbf{k}''$ ) beams at a plane interface between nonmagnetic insulators. The index of refraction of the lower medium (2) exceeds that of medium (1),  $n_2 > n_1$ . In this case the refracted beam is bent towards the normal; the electric field is polarized in the  $z$ -direction and the  $\mathbf{k}$  vectors lie in the  $x - y$  plane. (Problem 5.16)

Using the development of Chap. 4 [see, in particular, Problem 4.10], we see that it is the normal component of  $\mathbf{D}$  rather than of  $\mathbf{E}$  which is continuous across the boundary. Labeling the  $\mathbf{B}$  fields of the incoming and reflected beams in Fig. 5.8  $\mathbf{B}, \mathbf{B}',$ , respectively, and of the transmitted beam  $\mathbf{B}''$ , with similar superscripts for the electric fields  $\mathbf{E}$ , the boundary conditions on the fields at the surfaces can be written

$$\mathbf{B} + \mathbf{B}' = \mathbf{B}'' \quad (5.26a)$$

$$E_x + E'_x = E''_x \quad (5.26b)$$

$$E_z + E'_z = E''_z \quad (5.26c)$$

$$D_y + D'_y = D''_y. \quad (5.26d)$$

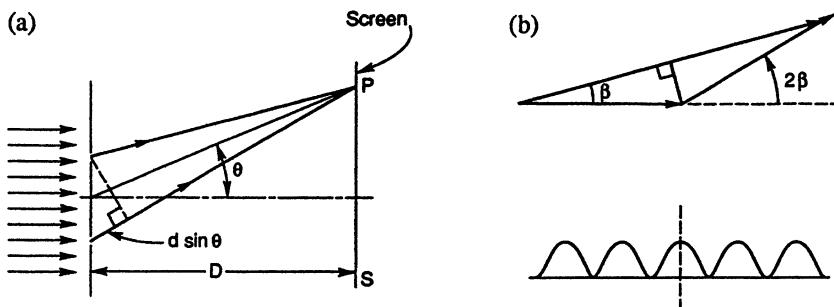
**Example 5.1.** A beam of plane coherent light is passed through two long slits separated a distance  $d$ . By considering each slit as a single Fresnel zone, i.e., by neglecting diffraction effects, show that the illumination of a screen a large distance  $D$  away will be given by

$$I = I_0 \cos^2 \beta$$

where  $I_0$  is the illumination at the center of the pattern, and

$$\beta = \pi d(\sin \theta)/\lambda$$

with  $\theta$  as in the diagram in Fig. 5.9a.



**Fig. 5.9.** (a) Light from two slits leads to interference at screen  $S$ , Example 5.1. (b) Example 5.1. Top, phasor diagram; bottom interference pattern with diffraction effects omitted

Draw the pattern, and give a direct physical explanation for the occurrence of minima at  $\beta = (\pi/2)(2n + 1)$   $n = 0, 1, 2, \dots$

*Solution.* With diffraction effects left out, an EM field at the point  $P$  in Fig. 5.9a can be represented by the phasor sum of the two individual fields, likewise represented by phasors. The component phasors are equal in length corresponding to equal amplitudes for the two fields, but are shifted with respect to each other by their difference in phase,  $2\beta$ :

$$2\beta = (2\pi/\lambda) \text{ (difference in optical paths} = d \sin \theta) .$$

Simple geometry (Fig. 5.9b) gives

$$|\text{resulting phasor}| = 2 \cos \beta |\text{each individual phasor}| ,$$

and  $I = I_0 \cos^2 \beta$  follows when we recall that intensities are proportional to the square of the field amplitudes. The minima in Fig. 5.9b occur whenever the two components point in opposite directions (optical paths differing by an integral number of half wavelengths). In practice the interference pattern is modulated by diffraction effects (Problems 5.10, 5.11).

**Example 5.2.** Let us prove that the angle of reflection  $\theta'$  is equal to the angle of incidence  $\theta$  relative to the  $x - z$  plane (Fig. 5.8) but for an arbitrary plane of incidence.

*Solution.* In time and space an electric field component of the incoming beam will satisfy at the surface

$$E_x(\mathbf{r}, t) = E_{x0} e^{i(k_x x + k_z z - \omega t)}. \quad (5.27)$$

The boundary condition (5.26b) thus becomes

$$E_{x0} e^{i(k_x x + k_z z - \omega t)} + E'_{x0} e^{i(k'_x x + k'_z z - \omega t)} = E''_{x0} e^{i(k''_x x + k''_z z - \omega t)} \quad (5.28a)$$

at every instant of time  $t$ , and at every point  $x, z$  on the infinite plane separating the two media. This condition can only be satisfied if, in the first place, the two fields oscillate together in time, i.e., the same frequency characterizes all the fields, as in fact we have already assumed in (5.28a). By the same token we must have

$$k_x = k'_x = k''_x \quad (5.28b,c)$$

and

$$k_z = k'_z = k''_z. \quad (5.28d,e)$$

On the other hand, the  $\mathbf{E}$  fields for both the incoming and outgoing beams satisfy the wave equation (5.22a) with the index of refraction  $n_1 = \sqrt{\varepsilon_1(\omega)/\varepsilon_0}$ , requiring that

$$k_x^2 + k_y^2 + k_z^2 = k'_x^2 + k'_y^2 + k'_z^2. \quad (5.29)$$

In view of (5.28b,d) and (5.29) we conclude that

$$k'_y = -k_y \quad (5.30)$$

giving the desired result  $\theta = \theta'$ . The reader is referred to Problem 5.16 for details analyzing the refraction shown in Fig. 5.8.

## 5.5 The Magic Rule in Three Dimensions

The main purpose of this section is to put the theory of diffraction and interference on a first-principles' basis, i.e., to relate the developments in this chapter to the wave equation for the electric or magnetic field with boundary conditions.

The formal development outlined here depends, in distinction to Sect. 3.5, upon the full Green's theorem in vector calculus, which is an extension of Gauss's theorem. Let  $U(x, y, z)$  and  $V(x, y, z)$  be any two scalar fields of  $x, y, z$ . For these fields there holds the vector identity

$$\nabla \cdot (U \nabla V) = (\nabla U) \cdot (\nabla V) + U \nabla^2 V. \quad (5.31)$$

We can integrate both sides of this identity over any closed volume, and directly apply Gauss's theorem to the left side. This yields

$$\oint \mathbf{U} \nabla V \cdot d\mathbf{S} = \int \nabla U \nabla V dv + \int U \nabla^2 V dv , \quad (5.32)$$

where the left integral is over the surface area enclosing the arbitrary volume. We can get a second equation like (5.32) by interchanging  $U$  and  $V$ . Subtracting this second equation from (5.32) we obtain the desired Green's theorem:

$$\oint [U \nabla V - V \nabla U] \cdot d\mathbf{S} = \int [U \nabla^2 V - V \nabla^2 U] dv . \quad (5.33)$$

Now, suppose  $\psi(\mathbf{r}, t)$  is the scalar solution of a wave equation with a driving term harmonic in time

$$\nabla^2 \psi(\mathbf{r}, t) - (1/c^2) \partial^2 \psi(\mathbf{r}, t) / \partial t^2 = -\varrho(\mathbf{r}) e^{-i\omega t} \quad (5.34)$$

with  $\varrho(\mathbf{r}) e^{-i\omega t}$  the source of the field. We have the solution

$$\psi(\mathbf{r}, t) = u(\mathbf{r}) e^{-i\omega t} ,$$

where  $u(\mathbf{r})$  satisfies the *Helmholtz equation*:

$$\nabla^2 u(\mathbf{r}) + k^2 u(\mathbf{r}) = -\varrho(\mathbf{r}) , \quad (5.35)$$

with  $k^2 = (\omega/c)^2$ . The Green's function technique consists in solving this equation by first splitting the source  $\varrho(\mathbf{r})$  into a spatial distribution of point sources  $\delta(\mathbf{r} - \mathbf{r}_0)$ , and ultimately taking the solution to be an appropriate integral of the respective individual solutions corresponding to the point sources.

The individual solutions are the Green's functions  $G(\mathbf{r}|\mathbf{r}_0)$

$$\nabla^2 G(\mathbf{r}|\mathbf{r}_0) + k^2 G(\mathbf{r}|\mathbf{r}_0) = -\delta(\mathbf{r} - \mathbf{r}_0) . \quad (5.36)$$

We shall learn that the boundary conditions of  $G(\mathbf{r}|\mathbf{r}_0)$  will need to be chosen in particular ways if  $G(\mathbf{r}|\mathbf{r}_0)$  is to be instrumental in solving the physical problem for  $\psi(\mathbf{r}, t)$ . We recall that the problem can be of two categories. The first category is as posed by the equation of motion (5.34) for  $\psi(\mathbf{r}, t)$ , that is, an inhomogeneous equation with certain fixed boundary conditions. A second category of problem is where the right (driving) side vanishes throughout the volume of interest, the source of the field being only in the boundary, for example, a cavity driven by currents in one of its walls. Even then, just as in Chap. 3,  $G(\mathbf{r}|\mathbf{r}_0)$ , which is a function satisfying the *inhomogeneous* equation (5.36), can be used to solve this second problem.

Although we are in this section singling out the Helmholtz equation, which originates with time harmonic forces in the wave equation, the Green's function concept carries over to field sources with arbitrary time dependence as, in fact, we saw in Chap. 2. At the same time the wave equation need not have the standard form of (5.34), for example, the Poisson equation  $\nabla^2 \phi = -\varrho/\epsilon_0$  of electrostatics can also be solved by a Green's function, as we did in Chap. 4.

In wave mechanics, somewhat related Green's functions play an important role.

Let us return to the Helmholtz Green's function  $G(\mathbf{r}|\mathbf{r}_0)$  [strictly  $G_k(\mathbf{r}|\mathbf{r}_0)$ ,  $k = \omega/c$ ]. Let us replace  $\mathbf{r}$  by  $\mathbf{r}_0$  in (5.35) and multiply both sides by  $G(\mathbf{r}|\mathbf{r}_0)$ ; multiply both sides of (5.36) by  $u(\mathbf{r}_0)$ , and subtract. With  $G(\mathbf{r}/\mathbf{r}_0) = G(\mathbf{r}_0/\mathbf{r})$  we get

$$G(\mathbf{r}|\mathbf{r}_0)\nabla_0^2 u(\mathbf{r}_0) - u(\mathbf{r}_0)\nabla_0^2 G(\mathbf{r}|\mathbf{r}_0) = u(\mathbf{r}_0)\delta(\mathbf{r} - \mathbf{r}_0) - G(\mathbf{r}|\mathbf{r}_0)\varrho(\mathbf{r}_0).$$

Now integrate both sides over an arbitrary volume of  $\mathbf{r}_0$  and take the integral of the term containing  $\varrho$  to the left side. We get

$$\begin{aligned} & \int \int \int [G(\mathbf{r}|\mathbf{r}_0)\nabla_0^2 u(\mathbf{r}_0) - u(\mathbf{r}_0)\nabla_0^2 G(\mathbf{r}|\mathbf{r}_0)] dv_0 \\ & + \int \int \int \varrho(\mathbf{r}_0)G(\mathbf{r}|\mathbf{r}_0) dv_0 = u(\mathbf{r}) \quad \text{if } \mathbf{r} \text{ is in the volume} \\ & \qquad \qquad \qquad = 0 \quad \text{otherwise}. \end{aligned}$$

Lastly, we can apply Green's theorem here, (5.33), to the first volume integral and write

$$\begin{aligned} \frac{u(r)}{0} &= \oint [G(\mathbf{r}|\mathbf{r}_0)\nabla_0 u(\mathbf{r}_0) - u(\mathbf{r}_0)\nabla_0 G(\mathbf{r}|\mathbf{r}_0)] \cdot d\mathbf{S}_0 \\ &+ \int \int \int \varrho(\mathbf{r}_0)G(\mathbf{r}|\mathbf{r}_0) dv_0, \end{aligned} \quad (5.37)$$

where in the first integral on the right  $\mathbf{r}_0$  lies on the surface  $A_0$  of the volume, while in the second  $\mathbf{r}_0$  lies anywhere inside.

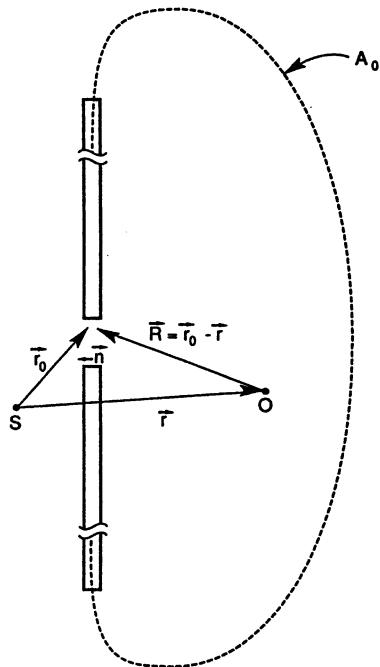
Equation (5.37) is, again, the magic rule, now in three dimensions. If we wish to do a problem in which  $u(\mathbf{r}_0)$  is specified at the boundary but  $\nabla_0 u(\mathbf{r}_0)$  is unknown there, we choose  $G(\mathbf{r}|\mathbf{r}_0)$  to vanish at the boundary. We can also handle the problem when it is the gradient of  $u(\mathbf{r}_0)$  at the surface which is given. (How?) Also, a solution is forthcoming regardless of whether  $\varrho(\mathbf{r}_0)$  exists or vanishes everywhere.

Let us apply (5.37) to finding, say, one component (polarization) of the electric field  $E(\mathbf{r})$  when light from a point source S is diffracted by a single aperture in an infinite baffle, Fig. 5.10. For the spherical wave emitted at S, which we regard as the origin of our system,  $\mathbf{r}_0 = 0$ , we write

$$E(\mathbf{r}_0) = E_0 e^{ikr_0}/r_0, \quad (5.38)$$

and identify  $E(\mathbf{r}_0)$  in (5.38) with  $u(\mathbf{r}_0)$  in (5.37), and  $E(\mathbf{r})$ , the desired disturbance at the point of observation O with  $u(\mathbf{r})$ . The surface of integration  $A_0$  is comprised of the dashed line in the figure, completed by the baffle and the aperture. The appropriate choice for  $G(\mathbf{r}|\mathbf{r}_0)$  is that of an outgoing wave from  $\mathbf{r}$ :

$$G(\mathbf{r}|\mathbf{r}_0) = \frac{e^{ik|\mathbf{r}_0 - \mathbf{r}|}}{|\mathbf{r}_0 - \mathbf{r}|}. \quad (5.39)$$



**Fig. 5.10.** Light from a point source  $S$ , diffracted by a single slit, gives rise to an electric field component  $E(\mathbf{r})$  at the point of observation,  $O$ . The dashed portion of the area  $A_0$  at infinity is completed by the baffle and aperture.  $\mathbf{n}$  is a unit vector, outward bound from the aperture

The same considerations determine  $G(\mathbf{r}|\mathbf{r}_0)$  in (5.39) as  $E(\mathbf{r}_0)$  in (5.38), both expressions satisfying wave equations with appropriate singularities on their right sides. We notice that  $G(\mathbf{r}|\mathbf{r}_0) = G(\mathbf{r}_0|\mathbf{r})$ , i.e. the *principle of reciprocity* is satisfied; it always will be for  $G$  (see Sect. 3.5). The dashed portion of the area  $A_0$  is at infinity, and of the two possible conditions there, outgoing or incoming waves, the first alternative needs to be used.  $u(\mathbf{r}_0)$  and  $\nabla_0 u(r_0)$  then vanish sufficiently rapidly that no integration need be carried out here; there are no reflections. The physical situation is further specified by saying that the surface integral in (5.37) vanishes over the baffle; that is,  $u(\mathbf{r}_0)$  and its gradient vanish there, whereas within the aperture,  $u(\mathbf{r}_0)$  and its gradient may be obtained from (5.38) – no reflections are coming back into the opening from anywhere. From (5.38), then

$$\mathbf{n} \cdot \nabla_0 u(\mathbf{r}_0) = -E(\mathbf{r}_0) \left( \frac{1}{r_0} - ik \right) \frac{\mathbf{r}_0}{r_0} \cdot \mathbf{n}, \quad (5.40a)$$

and, from (5.39)

$$\mathbf{n} \cdot \nabla_0 G(\mathbf{r}|\mathbf{r}_0) = G(\mathbf{r}|\mathbf{r}_0) \left( \frac{1}{|\mathbf{r} - \mathbf{r}_0|} - ik \right) \frac{\mathbf{r} - \mathbf{r}_0}{|\mathbf{r} - \mathbf{r}_0|} \cdot \mathbf{n}. \quad (5.40b)$$

Since the wavelength of the light,  $2\pi/k$ , is assumed to be small compared to distances  $r_0$  or  $|\mathbf{r}_0 - \mathbf{r}|$ , only the second terms in the parentheses of (5.40a,b) are retained.  $\mathbf{n}$  being the outward unit vector from the baffle, as shown in

Fig. 5.10, we have, substituting (5.38–5.40) into (5.37),

$$\begin{aligned} E(\mathbf{r}) &= ik \int_{\text{aperture}} \left\{ E(\mathbf{r}_0) G(\mathbf{r}|\mathbf{r}_0) \left[ \frac{\mathbf{n} \cdot \mathbf{r}_0}{r_0} - \frac{(\mathbf{r}_0 - \mathbf{r}) \cdot \mathbf{n}}{|\mathbf{r}_0 - \mathbf{r}|} \right] \right\} dS_0 \\ &= ikE_0 \int_{\text{aperture}} \frac{e^{ik(r_0 + |\mathbf{r}_0 - \mathbf{r}|)}}{(r_0)|\mathbf{r}_0 - \mathbf{r}|} \left[ \frac{\mathbf{n} \cdot \mathbf{r}_0}{r_0} - \frac{\mathbf{n} \cdot (\mathbf{r}_0 - \mathbf{r})}{|\mathbf{r}_0 - \mathbf{r}|} \right] dS_0. \end{aligned} \quad (5.41)$$

In case the source instead of being a point at  $S$  close to the screen is removed far away from the screen giving an incoming plane wave  $E_0 e^{ik \cdot \mathbf{r}_0}$ ,  $E(\mathbf{r})$  becomes

$$E(\mathbf{r}) = -ikE_0 \int_{\text{aperture}} \frac{e^{ik \cdot \mathbf{r}_0} e^{ik|\mathbf{r}_0 - \mathbf{r}|}}{|\mathbf{r}_0 - \mathbf{r}|} \left[ \frac{\mathbf{n} \cdot (\mathbf{r}_0 - \mathbf{r})}{|\mathbf{r}_0 - \mathbf{r}|} + 1 \right] dS_0 \quad (5.42)$$

where  $\mathbf{r}_0$  is now understood to be pointing parallel to  $-\mathbf{n}$ , i.e., to the incoming direction  $\mathbf{k}$ , so that  $e^{ik \cdot \mathbf{r}_0}$  is a phase factor constant over the area of the aperture, and can be left out. The term in square brackets is the “obliquity factor” mentioned in Sect. 5.2. Clearly, for the usual deflections of a few degrees or less we may likewise consider this factor to be a constant. In this way (5.42) leads to (5.7a).

## Problems

**5.1 Einstein Photoelectric Effect.** Einstein's law regarding the emission of electrons from metals illuminated by light of frequency  $\nu$  is

$$U_e = h\nu - \phi.$$

a) If the electrons in a metal occupy all available energy states, one electron per state, up to a maximum energy called the Fermi energy, explain why  $U_e$  is the maximum kinetic energy of any emitted electron.

b) Write an equation for the threshold frequency  $\nu_c$  below which you do not expect any electrons to be emitted.

c) An electron-volt is the amount of energy acquired by a particle having the numerical charge of one electron,  $1.60 \times 10^{-19}$  coulombs, when it is accelerated through a potential difference of 1 volt. If the work function  $\phi$  of a metal is 1 eV, what is the threshold frequency  $\nu_c$ , and what is the corresponding wavelength,  $\lambda_c$ .

If  $\phi = 1.82$  eV as for sodium, what is  $\lambda_c$ ?

d) Explain how the intensity of the illumination will affect the number of emitted photoelectrons

- i) for  $\nu > \nu_c$
- ii) for  $\nu < \nu_c$ .

These observations were not understood before Einstein's explanation of photoemission.

**5.2** a) If the frequency of radiation from some particular source were precisely known, what knowledge would we have about the time of emission  $t$ ?

b) If the uncertainty in the energy of the upper level of an emitter is given by

$$E \approx 10^{-20} J,$$

about how many wavelengths  $\lambda$  is the coherence length, when the emitted radiation is at  $\lambda \approx 5000\text{\AA}$ ?

**5.3** If an atom is known to have been excited at time  $t = 0$ , then the probability that it is still excited at time  $t$  is given by  $e^{-\gamma t}$ , where  $1/\gamma$  is the lifetime of the excitation. Assume that, proportional to the amplitude of a damped harmonic oscillator, the radiated field  $B(t)$  depends on  $t$  as:

$$\begin{aligned} B(t) &= 0 & t < 0 \\ &= b_0 \sqrt{\omega_0} \cos \omega_0 t e^{-\gamma t/2} & t > 0. \end{aligned}$$

a) Show that  $B(\omega)$ , the usual Fourier component of  $B(t)$ , is given by

$$B(\omega) = \frac{\sqrt{\omega_0} b_0}{2\pi 2i} \left[ \frac{1}{(\omega - \omega_0 - i\gamma/2)} + \frac{1}{(\omega + \omega_0 - i\gamma/2)} \right].$$

Show also that  $U(\omega)$ , the  $\omega$  contribution to the energy  $U = \int U(\omega) d\omega$ , is

$$U(\omega) = B^*(\omega)B(\omega)c/\mu_0.$$

*Hint:* See (4.17c, 19) and Problem 1.14.

Finally, show that the *leading* component of  $U(\omega)$  is proportional to

$$\frac{\omega_0}{[(\omega - \omega_0)^2 + \gamma^2/4]}$$

This expression is the well known “Lorentzian lineshape” yielded by a full quantum mechanical treatment of a two-level system.

b) Devise some method for estimating  $\Delta\omega$  from part (a),  $\Delta t$  from  $B(t)$ , and compare your estimate for the product  $(\Delta\omega)(\Delta t)$  with (5.4).

**5.4** Show, with a diagram, the somewhat trivial development of how the Huygens principle can be used to predict the advancement of a plane EM wave.

**5.5** For what values of the angle  $\alpha$  do we get minima in the intensity of the light diffracted by a single long slit, given (5.9, 10)? Can you give a physical interpretation for your result by analyzing the incoming radiation fields as comprised of contributions from a series of long horizontal elements of area (Fresnel zones) comprising the slit?

*Hint:* Consider the radiation from two Fresnel zones at a time.

**5.6** What happens to the single slit diffraction pattern when  $\lambda \gg a$ , i.e. the width is small compared even to the wavelength of the diffracted light, when  $\lambda \ll a$ ? Give a simple physical characterization for each of these situations.

**5.7** Find an analytic expression for the diffraction pattern of a square aperture of area  $a^2$ , with  $a \approx \lambda$ . Make a qualitative drawing representing your calculated light intensity. (The diffraction pattern turns out to be a square grid.)

**5.8** As mentioned in the text, an important measuring device is the diffraction grating, an opaque screen with many ( $N$ ) parallel slits.

a) Using the Fresnel zone arguments, as in Example 5.1, show that the first minimum for a grating will occur at  $\theta = \lambda/(Nd)$ , where  $d$  is the distance between adjacent slits.

*Hint:* For a principal maximum all phasors are parallel, for the corresponding minimum the phasors form a closed  $N$ -sided polygon.

b) We can say that two colors, wavelengths  $\lambda$  and  $\lambda + \Delta\lambda$ , are resolved when the second minimum for  $\lambda$  falls at the first principal maximum (not counting the central maximum) for  $\lambda + \Delta\lambda$ ,  $\Delta\lambda$  taken as positive.

Show that  $(\Delta\lambda/\lambda) = (1/N)$ .

What can you conclude about the relationship between the number of slits  $N$  and the “resolving power” of a diffraction grating?

**5.9** a) When we shift our attention from the first principal maximum to the  $m$ -th maximum (not counting the central peak), the so-called *m-th order* spectrum, we find that

$$\Delta\lambda/\lambda = 1/mN .$$

Can you show this?

b) In a grating the sodium doublet (5890 and 5895.9 Å) is viewed in third order at  $1^\circ$  to the normal, and is barely resolved. Find (i) the grating spacing and (ii) the total width of the rulings.

**5.10** A plane coherent light beam impinges on the baffle B, Fig. 5.11.  $S_I$ ,  $S_{II}$ ,  $S_{III}$  are three imaginary screens assumed to have been placed in succession at almost the same distance behind B, one at a time. It is suggested that the reader copy Fig. 5.11 and then:

a) On the screen  $S_I$  draw the diffraction pattern corresponding to a single slit of width  $a$  at the center of B. State the position of the first minimum. The wavelength is  $\lambda$ .

b) On screen  $S_{II}$  draw the intensity pattern assuming there are two horizontal, vertically centered slits in B, separated by a distance  $d$ , with  $d \simeq (9/2)a$ . Do not neglect diffraction effects in your diagram.

c) On screen  $S_{III}$  draw the pattern for the same values of  $a$  and  $d$  as in part (b), but assume there are many slits. Be sure the distinction with the pattern on  $S_{II}$  is clearly, if only qualitatively, displayed.

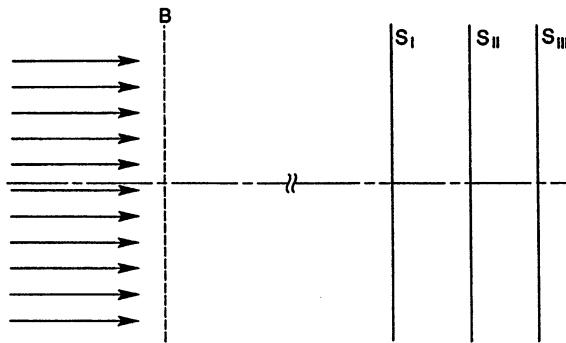


Fig. 5.11. Problem 5.10

**5.11** Consider a baffle with three slits, each of width  $a$ , separated by a distance  $d$ . Divide each slit into four zones. Write down a phase factor  $e^{ikR}$  for each of the twelve light rays joining the zones to a point at  $\theta$  on the infinitely wide screen,  $R$  being the distance from zone to screen. Show that the sum of the twelve corresponding phasors can be arranged into a product of a diffraction factor multiplied by an interference factor.

**5.12** At what distance will the two headlights of an oncoming car first appear as distinct? Assume that the determining limitation is the finite resolution of the eye, with a pupil diameter of about 1/2 cm, and that the illumination is at 5500 Å. The headlights are separated by 120 cm.

**5.13** Why does a Laue diffraction pattern display dots, whereas a Debye-Scherrer pattern display a series of concentric rings? How would you produce a monochromatic X-ray beam?

**5.14** a) Show that the (uncompensated) polarization charge per unit area at a boundary aligned perpendicular to a polarization  $\mathbf{P}$  is given by

$$\sigma_{\text{pol}} = P .$$

*Hint:* Assume  $n$  dipoles/unit area, each dipole having a moment of, say,  $qd$ .

b) Generalize this result to

$$\sigma_{\text{pol}} = \mathbf{P} \cdot \mathbf{n}$$

where  $\mathbf{n}$  is a unit vector perpendicular to the area now making an arbitrary angle with  $\mathbf{P}$ .

Demonstrate that the net charge *inside* a closed volume with polarization  $\mathbf{P}(\mathbf{r})$  is given by the surface integral

$$Q_{\text{pol}} \equiv \int \rho_{\text{pol}} dV = - \oint \mathbf{P} \cdot d\mathbf{A} .$$

c) Finally, use Gauss's theorem for vector fields to obtain the desired result

$$\rho_{\text{pol}} = -\nabla \cdot \mathbf{P} .$$

**5.15** Show that

$$\mathbf{D}(\mathbf{r}, t) = \mathbf{E}(\mathbf{r}, t) + \int_{-\infty}^{\infty} G(\tau) \mathbf{E}(\mathbf{r}, t - \tau) d\tau$$

$$\text{where } G(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} [\varepsilon(\omega) - 1] e^{-i\omega t} d\omega.$$

Reduce the result for  $\mathbf{D}(\mathbf{r}, t)$  when  $\varepsilon(\omega)$  is not a function of  $\omega$ . Note that, in general,  $\varepsilon(\omega)$  will be such as to ensure that  $G(t - \tau) = 0$ , when  $t < \tau$  (causality). Polarization can not precede an applied electric field.

**5.16** a) For the situation shown in Fig. 5.8, i.e., with the electric fields polarized to lie parallel to the  $z$ -axis, show that the amplitudes of the reflected and transmitted electric fields  $E'_0$  and  $E''_0$ , respectively, are related to  $E_0$ , the amplitude of the incoming beam by

$$E'_0 = \frac{(k_y - k''_y)}{(k_y + k''_y)} E_0$$

$$E''_0 = \frac{2k_y}{(k_y + k''_y)} E_0.$$

*Hint:* Use (5.26a) and (5.28b,c).

b) Prove the well-known Snell's law

$$n_2 \sin \theta'' = n_1 \sin \theta,$$

where  $\theta''$  is the angle of refraction, Figure 5.8, for both possible polarizations of the electric field.

**5.17** a) Show that a plane wave propagating in the  $z$ -direction with complex index of refraction  $n = n_R + in_I$  is given by

$$\mathbf{E}(z, t) = \mathbf{E}_0 e^{i\omega(n_R z/c - t)} e^{-\omega n_I z/c}.$$

b) Use the Maxwell equations to show that the wave equation in a non-polarizable, nonmagnetic metal of electric conductivity  $\sigma$  is given by

$$\nabla^2 \mathbf{E} - (1/c^2) \partial^2 \mathbf{E} / \partial t^2 - \mu_0 \sigma \partial \mathbf{E} / \partial t = 0.$$

Show that

$$n_I \simeq \sqrt{\sigma/2\varepsilon_0 \omega} \quad \text{when } \sigma \gg \varepsilon_0 \omega$$

as is the case in good conductors at any available frequency. However, for  $\omega \gtrsim 10^{13}$  rad/s dynamic effects associated with electron collisions set in.

c) If the electric field in the metal is attenuated as  $e^{-z/\delta}$ , where  $\delta$  is called the "skin depth", estimate  $\delta$  numerically for a microwave frequency of  $10^{10}$  Hz, given that  $\sigma \simeq 5.76 \times 10^7$  Ohm $^{-1}$  m $^{-1}$ .

d) Show that there exists between the complex conductivity  $\sigma(\omega)$  and the complex dielectric constant  $\epsilon(\omega)$  the general relation

$$\sigma(\omega) = -i\omega[\epsilon(\omega) - \epsilon_0].$$

Note and discuss the implied  $90^\circ$  phase difference between  $\sigma(\omega)$  and the electrical susceptibility  $\chi(\omega)$ .

*Hint:*

$$\nabla \times \mathbf{B} = \mu_0 \left[ \frac{\partial \mathbf{P}}{\partial t} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right].$$

Any current is to be viewed as a polarization current.

## Further Reading

D. Halliday, R. Resnick: *Physics*, Part Two (Wiley, New York 1978)

Readers who have had no previous exposure to physical optics may well wish to consult this more basic account.

G.R. Fowles: *Introduction to Modern Optics* (Holt, Rinehart and Winston, New York 1967)

An intermediate level treatment which the reader may find helpful.

M. Born, E. Wolf: *Principles of Optics* (Cambridge University Press, Cambridge, U.K. and New York, NY, 7th ed. 1999)

The greatest treatise on the subject of all time.

## 6. Wave Mechanics

**Abstract.** The need for a wave equation, given the de Broglie relation for particles and Bohr's spatially periodic description of atomic orbitals, is noted. A derivation of Schroedinger's wave equation as motivated by the classical wave equations of Chaps. 3 and 4 follows. The three postulates of wave mechanics – that to every observable there corresponds an operator, that the only possible values which measurements can yield are the eigenvalues, and that average values can be predicted with the Schrödinger wave function – are given. The special role of the energy eigenfunctions as yielding stationary states, and of the coordinate eigenfunctions as yielding a physical interpretation of the wave function are described. There then follows the analysis of the propagation of a free particle in terms of a wave packet, which serves as a basis for illustrating the Heisenberg uncertainty principle. Section 6.4 explores the wave–particle duality and indeterminism, and ends up by showing that these do not upset our notions of an orderly world. Section 6.5 develops the necessary aspects of the quantum physics of the harmonic oscillator, and, subsequently, the Green's function technique for treating the driven motion. Parallels with the classical problem are drawn. Lastly, Sect. 6.6 outlines some of the wide range of phenomena that have been quantitatively interpreted by quantum mechanics. The part played by Green's functions in quantum field theory is pointed out. The relevant Green's functions for both harmonic oscillators and free particles are related to the previous development of these functions in ordinary (first) quantization. Green's functions are seen to provide a bridge from classical to ordinary quantum physics to field theory. Example 6.4 explores the basis of the nuclear magnetic imaging technique (NMR) with a two level model, whereas Problem 6.19, using a similar model, deals with the origin of the resonance response of electronic polarizability. Both these examples again bring out the correspondence with the occurrence of the two distinct motions, natural and driven, in classical mechanics.

### 6.1 Origin of Schrödinger's Wave Equation

We have seen in the previous chapter that there existed at the beginning of the twentieth century a certain crisis in optics. Einstein's explanation of the photo–electric effect initiated a *wave–particle duality* for EM radiation,

since light not only had its wave nature, but could also act as a particle, the photon. There was, further, the necessity of explaining black body radiation. The ideal model for a black body is a cavity. If we have an em cavity, such as the ones discussed in Chap. 4, one can measure the intensity of radiation escaping from a little hole in any one of the walls as a function of the frequency of radiation  $\nu$ . Seen from the outside the hole is perfectly absorbing, black, and the emitted radiation will be exactly that of any black body. If the temperature of the cavity walls is carefully maintained at a value  $T$ , then according to classical statistical mechanics each normal mode in the cavity will radiate a certain energy depending only on its normal frequency  $\nu$ , and on  $T$ . (Statistical mechanics is the microscopic theory used to explain thermodynamics.) However, the total predicted radiation was far in excess of what was observed; in fact, it was divergent with the highest frequency  $\nu$ . As we mentioned in Chap. 5, Planck established a new empirical law giving the mode radiation as a function of  $\nu$  and  $T$ . As stated by Planck, the new law contained the physical constant  $h$ , Planck's constant.

The Austrian physicist Erwin Schrödinger (1887–1961) remarked that “wave mechanics was born in statistics” (statistical mechanics). The honor of first correctly proposing a wave-like nature for material particles belongs to the French physicist Louis de Broglie (1892–1987). In 1922, de Broglie was able to obtain an approximate formula for Planck's radiation law directly from statistical mechanics using only Einstein's  $E = h\nu$  for photons. In analyzing the radiation in the cavity as a gas of photons, de Broglie was continually reminded of traditional gas kinetics and statistical mechanics. “All of a sudden”, as he later said, he saw that the crisis in optics was due to the failure in understanding the true universal duality of waves and particles. Some ideas of his own on the “harmony of phases” then led him to extend to all material particles a law already suspected for photons. The famous de Broglie relation reads

$$\boxed{\lambda = h/p}, \quad (6.1)$$

where  $\lambda$  is the wavelength of the material particle or photon,  $h$  is again Planck's constant, and  $p$  the momentum of the particle.

When EM radiation is absorbed, the electric field exerts a force on the charges in the absorber. At the time of de Broglie's research, one knew that according to classical EM theory, energy absorption of  $U$  leads to a transfer of momentum  $U/c$  to the absorber. Thus, one suspected that for a photon, where  $U = h\nu$ ,  $p = h\nu/c = h/\lambda$ , as in (6.1). In 1923 the American physicist A.H. Compton verified this relationship for photons by scattering X-rays from the electrons of a graphite target. He noticed that scattered photons underwent a shift in wavelength commensurate with their shift in momentum.

Significant for our purpose is that (6.1) was also confirmed for electrons in 1927, by Clinton Davisson and Lester Germer in the United States and,

independently, by Sir George Thomson in Britain. The experiments involved diffraction of electron beams by crystals using the same theory as given for X-ray diffraction at the end of the previous chapter. Today, electron microscopes, which are a direct application of the de Broglie relation (see Problems), have become a widely used scientific instrument.

The de Broglie relation for material particles gave rise to the search for a wave equation. The idea that there exists an underlying wave nature for matter was strengthened by the notion, derived from atomic physics, that the electron orbitals in atoms were in some way spatially periodic. The work of the Danish physicist Niels Bohr (1885–1962) had shown that the angular momentum of the electrons was “quantized” into integer multiples of  $\hbar/2\pi$ . The discovery of a wave equation could lead to an understanding of phenomena on an atomic and possibly subatomic level, which, as was becoming increasingly evident, was not being provided by the physics known in 1924.

On the basis of what we have learned we might expect an underlying *wave function*  $\psi(x, t)$  dependent (restricting ourselves temporarily to one dimension) on  $x$ , and on the time  $t$ . This function might be a harmonic function, say, of the form

$$\psi(x, t) = u(x)v(t) , \quad (6.2a)$$

$$u(x) = \sin kx , \quad (6.2b)$$

$$v(t) = \sin \omega t . \quad (6.2c)$$

$k$ , the wave number  $2\pi/\lambda$ , is related through (6.1) to the momentum  $p$  by

$$k = p/\hbar , \quad (6.2d)$$

and  $\omega$  related to the total energy  $E$ , through the Einstein relation by

$$\omega = E/\hbar . \quad (6.2e)$$

Here  $\hbar$  is  $h/2\pi$ , i.e., Planck's constant divided by  $2\pi$  as in Chap. 5.

The spatial part of the wave equations which we encountered in Chaps. 3 and 4 read in one dimension

$$\frac{d^2u(x)}{dx^2} + k^2u(x) = 0 .$$

If one substitutes  $p/\hbar$  for  $k$  as in (6.2d), and recognizes that

$$\frac{p^2}{2m} = T \quad (6.3a)$$

$$= E - V(x) \quad (6.3b)$$

where  $m$  is the mass of the particle,  $T$  the kinetic energy,  $E$  the total energy and  $V(x)$  the potential energy, the latter taken as time-independent, one obtains

$$\left[ \frac{-\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \right] u(x) = Eu(x) . \quad (6.3c)$$

Equation (6.3c) is the time-independent Schroedinger equation, as written down by its great discoverer in early 1925. Generalizing to three dimensions, and taking  $V(r)$  to be the Coulomb potential of a hydrogen nucleus, Schroedinger was immediately able to use the equation to calculate the energies of the stationary atomic orbitals in the hydrogen atom. These energies then checked against observed energy differences deduced from the light spectrum of atomic hydrogen gas.

The full wave equation for a stretched string is

$$\frac{\partial^2 y}{\partial x^2} = \frac{1}{s^2} \frac{\partial^2 y}{\partial t^2}. \quad (3.3)$$

When we turn to the right side, i.e., now consider  $\partial^2 \psi / \partial t^2$ , we encounter a major difficulty. The result of taking two time derivatives of  $v(t)$  along with Einstein's equation (6.2e) gives rise to the square of the total energy, i.e., to  $E^2 \psi$ , rather than the  $E \psi$  we are seeking to balance against the left side of the equation. Indeed, the right side can be presumed to contain only a first derivative in time. This raises a different problem. If we are to have a solution to a first order differential equation, that solution will have to be an exponential function rather than a trigonometric one. Moreover, to avoid having the solution go to infinity or be exponentially damped as  $t$  goes to either plus or minus infinity but rather to get waves, the exponent in the solution will have to be imaginary. As a result the final *time-dependent Schroedinger equation* reads

$$\left[ -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \right] \psi(x, t) = i\hbar \frac{\partial \psi(x, t)}{\partial t}.$$

(6.4)

If  $V(x)$  is set equal to zero in (6.4), then as the reader can easily check, we get the solution

$$\psi(x, t) = e^{i(kx - \omega t)} \quad (6.5a)$$

with

$$\hbar\omega = \hbar^2 k^2 / 2m. \quad (6.5b)$$

Additionally, applying (6.2d) and (6.2e), we get

$$E = p^2 / 2m, \quad (6.5c)$$

the required result for a free particle.

For the general problem with a nonvanishing potential  $V(x)$ , (6.4) should be viewed as more general than (6.3c). The latter can be obtained from (6.4) through the trial solution  $u(x)v(t)$  of (6.2a), as we see in the Problems. One finds that

$$v(t) = e^{-i\omega t} \quad (6.6)$$

$$\omega = E/\hbar, \quad (6.2e)$$

that is, an Einstein frequency for the total energy,  $E$ . Concurrently,  $u(x)$  is satisfying the time-independent Schroedinger equation (6.3c), which proved so immediately successful in predicting atomic structure. Indeed, the time-dependent Schroedinger equation, (6.4), is the correct nonrelativistic wave equation for matter.

Lastly, we note that all of our discussion in this chapter is being restricted to time-independent potentials  $V(x)$ , i.e., closed systems which have constant total energy. It is this restriction which facilitates separation of space and time variables. Equation (6.4) itself is also valid for time-dependent potentials if one inserts for  $V(x)$  the appropriate potential  $V(x, t)$ . However, solution becomes more difficult, usually requiring so-called perturbation methods.

## 6.2 Postulates of Wave Mechanics

What is the relation between the Schroedinger wave function  $\psi(x, t)$  of the previous section and physical phenomena? The answer was far from obvious in 1925, especially in view of the complex variable nature of  $\psi(x, t)$ . Elucidation came as a result of the efforts of many physicists working together, in particular, Niels Bohr in Copenhagen, Max Born (1882–1970), Pascual Jordan (1902–1980), and Werner Heisenberg (1901–1976) all three in Gottingen, P.A.M. Dirac (1902–1984) in Cambridge, and Schroedinger himself, then in Zurich. Schroedinger discovered the relationship between results obtained with his wave equation and those obtained with matrix calculations initiated by Heisenberg in the summer of 1925. (The term *wave mechanics* refers to the treatment based on Schroedinger's wave equation, in contrast to the equivalent *matrix mechanics*, the two methods comprising quantum mechanics.)

The postulates of wave mechanics are as follows:

1. *To every physical observable there corresponds an operator. Operators replace the simple functions of classical physics. The table below, suitable for three dimensional calculations for a particle of mass m, explains:*

Observable	Classical function	Quantum mechanical operator
Position in cartesian coordinates	$\mathbf{r} = xi + yj + zk$	$\mathbf{r} = xi + yj + zk$
Momentum	$m[(dx/dt)\mathbf{i} + (dy/dt)\mathbf{j} + (dz/dt)\mathbf{k}]$	$(\hbar/i)\nabla$
Kinetic energy	$p^2/2m$	$-\hbar^2\nabla^2/2m$
Total energy	$(p^2/2m) + V(\mathbf{r})$	$-\hbar^2\nabla^2/2m + V(\mathbf{r})$

2. The only possible values which a measurement of the observable whose operator is  $\mathbf{Q}$  can yield are the eigenvalues  $q_n, q_l$  of the equations:

$$\boxed{Q u_n(\mathbf{r}) = q_n u_n(\mathbf{r}), \quad n = 1, 2, 3, \dots} \quad (6.7a)$$

$$\boxed{Q u_l(\mathbf{r}) = q_l u_l(\mathbf{r}) \quad \infty > l > -\infty} \quad (6.7b)$$

Here the functions  $u_n(\mathbf{r}), u_l(\mathbf{r})$  must be finite, continuous and single-valued, and must satisfy boundary conditions appropriate to the physical problem. In writing (6.7a,b) we distinguish between functions with discrete eigenvalues, (6.7a), and continuous eigenvalues, (6.7b), i.e.,  $l$  is a continuous variable. (In practice, the operator  $Q$ , as well as the boundary conditions determine whether we are dealing with continuous or discrete eigenvalues, or both.) Further, the  $u_n(\mathbf{r}), u_l(\mathbf{r})$  must be normalizable. In one dimension, for the first type of *eigenfunction*,  $u_n(x)$ , the normalization is contained in

$$\boxed{\int_{-\infty}^{\infty} u_n^*(x) u_{n'}(x) dx = \delta_{nn'}} \quad (6.8a)$$

For the second type:

$$\boxed{\int_{-\infty}^{\infty} u_{l'}^*(x) u_l(x) dx = \delta(l' - l)} \quad (6.8b)$$

Here  $\delta_{nn'}$  is the Kronecker delta,  $\delta(l' - l)$  the Dirac delta function. All the  $u(x)$  belonging to any operator together will turn out to be an orthonormal basis. (The basis may contain functions with both discrete and continuous eigenvalues, or only one type.) For the moment we must illustrate what is meant by eigenfunctions and eigenvalues with a few examples.

Suppose that the operator  $\mathbf{Q}$  is the total energy operator,  $\mathbf{Q} = \mathbf{H}$ . It is customary to use  $\mathbf{H}$ , calling the total energy operator the *Hamiltonian* in honor of Sir W.R. Hamilton (1805–1865), who reformulated classical Newtonian mechanics in terms of variational principles. Equation (6.7) generalized to three dimensions reads

$$\mathbf{H} u_n(\mathbf{r}) = E_n u_n(\mathbf{r}). \quad (6.9)$$

We consider a problem already referred to, namely, the motion of an electron in a hydrogen atom,

$$\mathbf{H} = -\hbar^2 \nabla^2 / 2m - e^2 / 4\pi\epsilon_0 r, \quad (6.10)$$

where  $e$  is the charge of the electron.  $\mathbf{H}$  is seen to be the sum of the kinetic energy and the Coulomb potential of the nucleus, the latter is considered

fixed. Suppose we restrict ourselves to bound situations, where the hydrogen atom does not disintegrate into an electron and nucleus, i.e., for total energy  $E < 0$ . The reader will suppose that (6.9) can be solved for any value of  $E$ , and so it can. However, we have seen in Chap. 3 that the imposition of boundary conditions limits physical solutions of differential equations. The boundary conditions for  $u_n(r)$  are supplied by the normalization conditions (6.8). Evidently,  $u_n(r)$  must go to zero as  $r$  goes to infinity, and must do so as  $(1/r)$  or faster.

It turns out that for  $E < 0$  only certain discrete values of  $E, E_1, E_2, \dots$ , lead to solutions of (6.9) that satisfy the boundary conditions. These discrete values of  $E$  are the eigenvalues. They are precisely “the possible values of the total energy” of the postulate. (A discussion of what happens when  $E > 0$  can be found in Chap. 7). For the hydrogen atom with  $E < 0$  the eigenvalues are the quantum energy levels with which the reader may be familiar from elementary modern physics or from chemistry. On the other hand the functions which solve the partial differential equation obtained by combining (6.9) and (6.10) may not be familiar. Indeed, although he was a powerful mathematician in his own right, Schroedinger in 1924–25 had to consult a mathematician friend to fully solve the problem. (Nowadays physicists do learn all about the functions. See Problem 6.3 for the solution corresponding to the lowest energy level.) There is, however, an important problem which we are in a position to solve in detail, and we now do so as a second illustration of eigenfunctions and eigenvalues.

Consider a particle of mass  $m$  in a one dimensional potential, Fig. 6.1:

$$\begin{aligned} V(x) = 0 & \quad L \geq x \geq 0 \\ V(x) = \infty & \quad x > L, x < 0. \end{aligned} \quad (6.11)$$

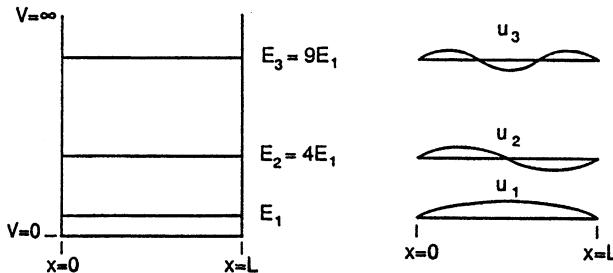
This problem is actually of great importance in modern physics, tending towards a simple model for nucleons bound inside the nucleus, or of electrons confined to a metallic crystal. It also is directly related to the model of photons confined to a cavity, as analyzed in the theory of black body radiation (based on the normal mode problem in a cavity as formulated in Chap. 4, see Sect. 6.5).

Inside the one dimensional box (6.9) takes the form

$$\frac{-\hbar^2}{2m} \frac{d^2 u_n(x)}{dx^2} = E_n u_n(x). \quad (6.12)$$

We note that we have chosen the functions  $u_n(x)$  to be time-independent. We may always do so when the operator in the eigenfunction equation (6.7) itself does not depend on the time. One can then always solve (6.7) by the method of separation of variables, i.e., through the trial solution (6.2a). Outside the box, where the potential is infinite, no solution to (6.3) is possible other than  $u_n(x) = 0$ . Thus, we take the boundary conditions on  $u_n(x)$  to be

$$u_n(L) = u_n(0) = 0. \quad (6.13)$$



**Fig. 6.1.** On the *left*, a potential well with infinitely high vertical walls. Also shown are the first three energy levels. On the *right*, the corresponding time-independent wave functions  $u_n(x)$

Apart from constants, (6.12) and (6.13) are identical to (3.9b) and (3.13), which describe the normal modes of a string of length  $L$  with fixed end points. We therefore have the familiar solutions

$$u_n(x) = \sqrt{2/L} \sin(n\pi x/L), \quad n = 1, 2, 3, \dots \quad (6.14)$$

where  $u_n(x)$  is already normalized, i.e., satisfies (6.8a). By substitution we obtain the energy eigenvalues:

$$E_n = \hbar^2 k_n^2 / 2m, \quad (6.15a)$$

$$k_n = n\pi/L. \quad (6.15b)$$

Again, the allowed energy levels are quantized.

Since the operator  $\mathbf{H}$  is the space-dependent operator in the Schroedinger time-dependent equation (6.4), finding the eigenfunctions and eigenvalues of  $\mathbf{H}$  also yields the general form of the solution of (6.4), namely,

$$\psi(x, t) = \sum_{n=1}^{\infty} A_n u_n(x) e^{-iE_n t/\hbar} \quad (6.16)$$

where the  $A_n$  are time-independent constants determined by the initial value  $\psi(x, 0)$ .

Again, the functions  $u_n(x)$  are examples of eigenfunctions. Finding the normal modes in classical mechanics is a special example of the more general mathematical procedure of finding eigenfunctions and eigenvalues.

The operator corresponding to physical quantities in quantum mechanics turn out to be *Hermitian* (after the French mathematician Charles Hermite, 1822–1901). Such operators are defined in Sect. 6.5. They have real eigenvalues, see Problem 6.14. Since the eigenvalues are possible observed values, real eigenvalues are obviously a necessity for a consistent theory.

Let us now allow  $L$ , the length of the one dimensional box of Fig. 6.1, to get large, i.e.,  $L \rightarrow \infty$ , and look for the eigenvalues of the momentum operator

$$\mathbf{p}_x = -i\hbar\partial/\partial x.$$

Thus, we are to solve the equations

$$-i\hbar \partial u(x)/\partial x = p_k u(x). \quad (6.17)$$

Again, the solutions are at hand, from Chap. 1. They are

$$u_k(x) = 1/\sqrt{2\pi} e^{ikx} \quad (6.18)$$

with

$$p_k = \hbar k. \quad (6.2d)$$

Equation (6.2d) is the expected relation between the momentum  $p$  and the wave number  $k$ , being again in accord with the de Broglie relation. If we suppose that  $L \rightarrow \infty$ , i.e., the boundaries do not affect the problem, then any value of  $k$  is allowed; the eigenvalues are continuous. We thus have from (6.8b)

$$\int_{-\infty}^{\infty} u_{k'}^*(x) u_k(x) dx = \delta(k' - k). \quad (6.19)$$

Actually, we see that we are dealing with the same functions  $u_k(x)$  already defined in (1.32a) of Chap. 1. The rigorous mathematical statement corresponding to (6.19) is

$$\int_{-\infty}^{\infty} u_{k'}^*(x) u_k(x) e^{-x^2/4n} dx = g_n(k - k'), \quad (1.32b)$$

where it is understood that the limit  $n \rightarrow \infty$  is of interest. In general, in physics we write and directly use the result as given in (6.19), but if we are not certain of whether a particular calculation is justifiable, we revert to the more rigorous formulation given in Chap. 1.

Lastly, we inquire into the eigenfunctions of the position variable  $x$ :

$$x u_\xi = \xi u_\xi(x). \quad (6.20)$$

Obviously the solution of this equation vanishes when  $x \neq \xi$ . P.A.M. Dirac discovered the need in quantum mechanics of the delta function. Using it, we have

$$u_\xi(x) = \delta(x - \xi) \quad (6.21)$$

with, again, the normalization (6.8b):

$$\int_{-\infty}^{\infty} \delta(x - \xi) \delta(x - \xi') dx = \delta(\xi - \xi'). \quad (6.22)$$

The reader may wish to consult Example 1.4 for the formulation of (6.20) and (6.21) in terms of the theory of generalized functions.

What happens to momentum eigenfunctions when  $L$  is finite? The answer, which we shall understand better shortly, is that there is no solution to (6.17) which satisfies the boundary conditions (6.13). In particular, the reader can check that the energy *eigenstates*  $u_m(x)$  of (6.14) do not satisfy (6.17). (The word “state” in eigenstate implies that the eigenfunction defines the particle’s situation and behavior – its state.) Thus, we now ask what happens if the Schroedinger wave function  $\psi(x, t)$  is not an eigenfunction of the operator for the observable in which we happen to be interested.

*3. When the particle has a wave function  $\psi(x, t)$ , the expected mean of a sequence of measurements of the observable whose operator is  $\mathbf{Q}$  is given by*

$$\bar{\mathbf{Q}}(t) = \int_{-\infty}^{\infty} \psi^*(x, t) \mathbf{Q} \psi(x, t) dx \quad (6.23a)$$

providing we normalize  $\psi^*(x, t)$  as

$$\int_{-\infty}^{\infty} \psi^*(x, t) \psi(x, t) dx = 1. \quad (6.23b)$$

The wording of this postulate makes us aware for the first time of a radical departure of quantum mechanics from classical mechanics. Classical mechanics sets out to, and does, fully predict future dynamics on the basis of present conditions, i.e., there is determinism. Quantum mechanics, on the other hand, in its utilization, is statistical in nature. This departure caused, as we shall see, wide discussion among physicists when quantum mechanics was first formulated, and even at the time of this writing, there is still some debate, especially in philosophy. However, in general, physicists accepted a standard formulation of quantum mechanics within a few years of its first formulation.

It is convenient in quantum mechanics to take over from statistics not only the idea of a mean as defined by (6.23a,b), but also a root square deviation:

$$\Delta Q \equiv \left[ (\mathbf{Q} - \bar{\mathbf{Q}})^2 \right]^{1/2} \quad (6.24a)$$

with

$$\begin{aligned} \overline{(\mathbf{Q} - \bar{\mathbf{Q}})^2} &= \int_{-\infty}^{\infty} \psi^*(x, t) [\mathbf{Q}^2 - 2\mathbf{Q}\bar{\mathbf{Q}} + (\bar{\mathbf{Q}})^2] \psi(x, t) dx \\ &= \overline{\mathbf{Q}^2} - (\bar{\mathbf{Q}})^2. \end{aligned} \quad (6.24b)$$

$\bar{\mathbf{Q}}$  being in one dimension a number rather than an operator, as in (6.23a),  $\mathbf{Q}^2$  meaning the operator  $\mathbf{Q}$  is applied twice in succession.

We recall the orthogonality of the various normal modes in the problems of previous chapters, for example, of the standing wave solutions for the stretched string in Chap. 3, and that these normal modes formed a basis. By “basis” we meant that functions of physical interest in the given interval, and satisfying the appropriate boundary conditions could be conveniently expressed in terms of these modes. It turns out that these properties of normal modes can be generalized; the eigenfunctions of a quantum mechanical operator belonging to different eigenvalues are orthogonal, and together form a basis, as already anticipated following (6.8a) and (6.8b). If more than one functionally independent eigenfunction belong to the same eigenvalue of an operator, as sometimes happens, an orthonormal basis can still be constructed.

The completeness of a set of eigenfunctions means that the Schroedinger wave function  $\psi(x, t)$  can always be expanded in them as a sum,

$$\boxed{\psi(x, t) = \sum_{n=1}^{\infty} B_n(t) u_n(x),} \quad (6.25a)$$

if we are dealing with a discrete basis, and as an integral,

$$\boxed{\psi(x, t) = \int_{-\infty}^{\infty} B_l(t) u_l(x) dl,} \quad (6.25b)$$

if we are dealing with a continuous basis. (In some problems both sum and integral are required.) The  $B(t)$  are time-dependent expansion coefficients. An important point to note is that if, and only if, the  $u$ 's are eigenfunctions of the total energy operator  $\mathbf{H}$ , will  $B(t)$  have the explicit form

$$B_n(t) = A_n e^{-iE_n t/\hbar}, \quad (6.26)$$

for  $n$  discrete, and a similar form for  $l$  continuous. Here  $A_n$  is independent of the time. This result was already indicated in (6.16).

Insertion of (6.25a) into the left of the normalization condition (6.23b) yields

$$\sum_{n'n} B_n^* B_n \int_{-\infty}^{\infty} u_{n'}^*(x) u_n(x) dx = \sum_n |B_n|^2$$

after using (6.8a), the orthonormality condition for the  $u_n$ 's. Thus, we find

$$\boxed{\sum_n |B_n(t)|^2 = 1.} \quad (6.27a)$$

Similarly with the continuous expansion (6.25b), we find for the left side of (6.23b) that

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} B_l'^* B_l \int_{-\infty}^{\infty} u_l'^*(x) u_l(x) dx dl dl' = \int_{-\infty}^{\infty} |B_l|^2 dl$$

with the help of (6.8b). One, therefore, also has

$$\int_{-\infty}^{\infty} |B_l(t)|^2 dl = 1. \quad (6.27b)$$

It becomes significant now to insert (6.25a) into (6.23) with the  $u_n(x)$  eigenfunctions of  $\mathbf{Q}$ :

$$\bar{\mathbf{Q}}(t) = \sum_{n,n'} B_{n'}^*(t) B_n(t) \int_{-\infty}^{\infty} u_{n'}^*(x) \mathbf{Q} u_n(x) dx.$$

Recalling (6.7a),

$$\mathbf{Q} u_n(x) = q_n u_n(x)$$

we find

$$\bar{\mathbf{Q}}(t) = \sum_n |B_n(t)|^2 q_n. \quad (6.28a)$$

Similarly, for functions with continuous eigenvalues,  $\mathbf{Q}(t)$  is given by the triple integral

$$\bar{\mathbf{Q}}(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} B_{l'}^*(t) B_l(t) \int_{-\infty}^{\infty} u_{l'}(x) \mathbf{Q} u_l(x) dx dl dl'$$

or

$$\bar{\mathbf{Q}}(t) = \int_{-\infty}^{\infty} q_l |B_l(t)|^2 dl. \quad (6.28b)$$

If (6.28a,b) is viewed in conjunction with (6.27a,b), the expansion coefficients  $B_n(t)$ ,  $B_l(t)$  in (6.25a,b) assume the physical interpretation that the squares of their moduli give the relative frequency of occurrence of the eigenvalues  $q_n$ ,  $q_l$ , respectively, at the time  $t$ .

Again, expansion in eigenfunctions of  $\mathbf{H}$  can be particularly useful. From (6.26) giving an exponential time-dependence for the expansion coefficient  $B_n(t)$ , we see that the square of the modulus  $|B_n(t)|^2$  is time-independent and, hence, the total energy given by (6.23a) will be constant, as expected for a closed system. (We use  $E$  for energy eigenvalues in place of  $h$  so as not

to confuse them with Planck's constant.) If it now happens that in such an expansion all the coefficients  $B_n(t)$  vanish except for one, say  $B_m(t)$ , then it is said that the particle is in the eigenstate  $u_m(x)$  and, additionally, it is said to be in a *stationary state*. Here  $\psi(x, t)$  assumes the simple form

$$\psi(x, t) = u_m(x)e^{-iE_m t/\hbar}. \quad (6.29)$$

The expectation value of the total energy  $E$  will then exactly equal  $E_m$ , and the spread in the total energy,  $\Delta E$ , defined in (6.24a), will vanish.

Moreover, from (6.23a) defining the expectation value of any observable, we see that when (6.29) is true, all expectation values  $\bar{Q}$  will be time-independent, as we might expect for a "stationary state" (see also Example 6.1).

Whether or not  $\psi(x, t)$  is in an energy eigenstate depends upon the initial conditions. The situation is, therefore, entirely analogous to that encountered in classical wave physics, where the distribution of the total energy among the normal modes was similarly determined at the outset, and then similarly stayed put.

The eigenstate with the lowest energy is referred to as the *ground state*, for example,  $u_1(x)$  for the square well problem, Fig. 6.1, or the stable 1s state of an electron in the hydrogen atom. The stability of this state comes simply from the fact that it takes a positive energy input from the outside to dislodge the particle from this state.

Let us now return to the solutions for a finite box of dimension  $L$  in (6.14–16). As has been mentioned, the stationary solutions

$$\psi_n(x, t) = u_n(x)e^{-E_n t/\hbar}, \quad (6.29)$$

$$u_n(x) = \sqrt{2/L} \sin(n\pi x/L), \quad n = 1, 2, 3 \dots \quad (6.14)$$

while eigenstates of the total energy operator  $H$ , are not eigenstates of the momentum operator  $p$ . Indeed, we can now calculate the expectation value  $\bar{p}$ , using (6.28a). It turns out (Example 6.1) that  $p = 0$  for any one of these states  $\psi_n(x, t)$ . The reason is that  $u_n(x)$  is a standing wave, i.e., composed of the two waves  $e^{ikx}/\sqrt{2L}, e^{-ikx}/\sqrt{2L}$ , each of which carries momentum, but the two momenta are pointing in opposite directions. The particle is moving back and forth between the two confining walls. We expect that  $\Delta p \neq 0$ , see Problems.

More is to be learned by looking at the two specific bases with continuous eigenvalues with which we are already familiar. First, suppose that we choose the position eigenfunctions  $u_\xi(x)$  of the operator  $x$  given by (6.21) as  $\delta(x - \xi)$ . With this choice of  $u_l(x)$  the integral expansion (6.25b) for  $\psi(x, t)$  reads

$$\psi(x, t) = \int_{-\infty}^{\infty} B_\xi(t) \delta(x - \xi) d\xi \quad (6.30a)$$

which means that here

$$B_\infty(t) = \psi(\infty, t). \quad (6.30b)$$

It follows from (6.28b) for any  $\bar{Q}$  that

$$\bar{x}(t) = \int_{-\infty}^{\infty} \psi^*(x, t)x\psi(x, t)dx, \quad (6.31)$$

with (6.30b) for the coefficients  $B_l(t)$ . Similarly, from (6.27b) it follows that

$$\int_{-\infty}^{\infty} \psi^*(x, t)\psi(x, t)dx = 1. \quad (6.23b)$$

One has thus recaptured the normalization condition on the Schroedinger wave function  $\psi(x, t)$ , thereby justifying the delta function normalization on the continuous eigenfunctions  $\delta(x - \xi)$ . Going now to our interpretation of  $|B_l(t)|^2$  as the relative frequency of occurrence of a continuous eigenvalue  $l$ , we obtain from (6.30b) a physical interpretation for  $\psi^*(x, t)\psi(x, t)$ :

$\psi^*(x, t)\psi(x, t)$  is the probability of finding the particle at  $x$  at time  $t$ ,

or  $\psi^*(x, t)\psi(x, t)dx$  is the probability of the particle's being between  $x$  and  $x + dx$  at  $t$ . One can deduce the same physical interpretation, less formally, directly from (6.31) for  $\bar{x}(t)$ , with the normalization condition (6.23b) signifying that the particle must exist somewhere; the sum of all the probabilities of finding the particle is unity.

Lastly, suppose that for  $u_l(x)$  we choose the momentum eigenfunctions  $u_k(x)$  of (6.18), i.e.,  $e^{ikx}/\sqrt{2\pi}$ . In that case the expansion (6.25b) becomes

$$\psi(x, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} B_k(t)e^{ikx}dk, \quad (6.32)$$

the standard Fourier representation (1.24) familiar to us from Chap. 1. It is thus an easy matter given  $\psi(x, t)$  to find  $B_k(t)$ , since we already know how to invert Fourier transforms.

We recall that the eigenvalues of the  $p$  operator,  $(\hbar/i)\partial/\partial x$ , are just  $\hbar k$ . The general formula for expectation values with continuous eigenvalues (6.28b) assumes the particularly transparent form

$$\bar{p}_x(t) = \int_{-\infty}^{\infty} \hbar k |B_k(t)|^2 dk \quad (6.33a)$$

when we expand in momentum eigenfunctions. Analogously combining (6.30b) and (6.31) we can write:

$$\bar{x}(t) = \int_{-\infty}^{\infty} x |B_x(t)|^2 dx . \quad (6.33b)$$

We notice that the  $u_k(x)$  are, for free particle, eigenfunctions of both the total energy operator  $\mathbf{H}$ , [i.e., the kinetic energy operator  $(-\hbar^2/2m)\partial^2/\partial x^2$  since  $V(x) = 0$ ], as well as of the momentum operator  $-i\hbar\partial/\partial x$ . Accordingly, it is tempting to employ a single wave  $u_k e^{-iE_k t/\hbar}$ ,  $E_k$  being  $\hbar^2 k^2 / 2m$ , as the solution of the Schrödinger wave equation  $\psi(x, t)$  in problems where the particle's momentum is well defined,  $\mathbf{p} = \hbar k$ ,  $\Delta p = 0$ . However, in view of the delta function normalization (6.8b) for the  $u_k(x)$ , one can not do so directly. Instead one must employ a stratagem like saying that one knows that the particle exists with a certain probability (intensity if there is a *beam* of particles, as there usually is) between certain points  $a < x < b$ , and then normalizing  $\psi(x, t)$  in that range to the appropriate overall probability for the range. Another common stratagem is to employ so-called *periodic boundary conditions*, an approach outlined in one of the problems. On the other hand, the delta function  $\delta(x - \xi)$ , which has the fundamental property

$$\int_{-\infty}^{\infty} \delta(x - \xi) dx = 1 \quad (1.19b)$$

[not to be confused with the orthonormalization condition (6.22)] becomes very useful when we wish to localize a particle, i.e., its probability density  $\psi^*(x, t)\psi(x, t)$ . These comments, pertaining to practice, will become clearer in the next section.

## 6.3 Motion of a Free Particle. The Heisenberg Uncertainty Principle

In the previous section we gave the formal postulates of wave mechanics, and began to solve some problems. To get a better feeling for how wave mechanics works, we shall solve the most basic problem of all in classical wave mechanics, that of a free particle, which according to Newton's first law, if not at rest, "continues in a state of uniform motion". How do we extract the equivalent of  $x = vt$  from Schrödinger's equation and the postulates of wave mechanics?

The procedure is to solve an initial value problem, using the Schrödinger equation as the governing wave equation, much as we solved the diffusion equation in Problem 1.17, or the problem of a traveling pulse on a stretched string in Chap. 3 using the wave equation there.

We shall write for our initial condition

$$\psi(x, 0) = \alpha e^{(-x^2/4a^2) + ik_0 x} . \quad (6.34a)$$

Here the proportionality constant  $\alpha$  is determined by the normalization condition:

$$\int_{-\infty}^{\infty} \psi^*(x, 0)\psi(x, 0) dx = 1 . \quad (6.23b)$$

We remember that (6.23b) ensures that the total probability of finding the particle *somewhere* is unity. We have

$$\alpha^2 \int_{-\infty}^{\infty} e^{-x^2/2a^2} dx = \alpha^2 \sqrt{2a^2\pi} = 1 ,$$

where we have used (1.5) to get our integral. Thus, normalized,  $\psi(x, 0)$  is

$$\psi(x, 0) = (2\pi a^2)^{-1/4} e^{(-x^2/4a^2)+ik_0x} . \quad (6.34b)$$

Why do we choose such a complicated starting function? Why do we not choose the initial probability distribution  $\psi^*(x, 0)\psi(x, 0)$  to be  $\delta(x)$ ? At least we will have nailed down the starting position of the particle to be at  $x = 0$ . These questions bring us to the heart of wave mechanics. To give a full answer we are obliged to Fourier-represent  $\psi(x, 0)$  in (6.34b) [see (1.24)]:

$$\psi(x, 0) = 1/\sqrt{2\pi} \int_{-\infty}^{\infty} A(k) e^{ikx} dk \quad (6.35a)$$

with

$$A(k) = 1/\sqrt{2\pi} \int_{-\infty}^{\infty} \psi(x, 0) e^{-ikx} dx \quad (1.33)$$

$$= (2a^2/\pi)^{1/4} e^{-(k-k_0)^2 a^2} . \quad (6.35b)$$

Equation (6.35b) follows upon application of (1.18), namely

$$\int_{-\infty}^{\infty} e^{ikx} e^{-ax^2} dx = \sqrt{\pi/a} e^{-k^2/4a} . \quad (1.18)$$

With (6.35b) for  $A(k)$ , we can use (6.33a) of the previous section to find the average momentum as

$$\bar{p}_x(0) = \hbar \int_{-\infty}^{\infty} k |A(k)|^2 dk . \quad (6.36a)$$

Now, the reader can easily use (1.18) to check that  $\int_{-\infty}^{\infty} |A(k)|^2 dk$  is equal to one. Indeed, we know from (6.23b) and (6.27b) for  $t = 0$  that this integral will be unity. We, therefore, find at once that

$$\begin{aligned} \int_{-\infty}^{\infty} k |A(k)|^2 dk &= \int_{-\infty}^{\infty} (k - k_0) |A(k)|^2 dk + k_0 \int_{-\infty}^{\infty} |A(k)|^2 dk , \\ &= k_0 \end{aligned}$$

since the first integral on the right, having an odd integrand, vanishes.

Therefore, from (6.36a)

$$\bar{p}(0) = \hbar k_0 , \quad (6.36b)$$

where here and subsequently we drop the subscript  $x$ .

Equation (6.36b) explains the factor  $e^{ik_0x}$  in our chosen initial condition. Without it the average momentum would be zero, and we shall find that then the particle would on the average not move at all.

We can also use (6.35b) for  $A(k)$  to find the spread in the momentum  $\Delta p$  defined as in (6.24a,b). In the Problems one obtains

$$\Delta p(0) = \hbar/2a . \quad (6.37)$$

At the same time the spread in  $x$ , denoted  $\Delta x$ , associated with the initial condition (6.34b) for  $\psi(x, 0)$  is found in Problem 6.6 to be

$$\Delta x(0) = a . \quad (6.38)$$

This calculated result for  $\Delta x$  comes as no surprise. The initial probability distribution is from (6.34b):

$$\psi^*(x, 0)\psi(x, 0) = (2\pi a^2)^{-1/2} e^{-x^2/2a^2} ; \quad (6.39)$$

the position of the particle can, indeed, be seen distributed approximately over an extension  $\Delta x \sim a$ . Further, if we replace  $2a^2$  by  $1/n$ , we formally get back from (6.39) for  $|\psi(x, 0)|^2$  the function

$$g_n(x) = \sqrt{n/\pi} e^{-nx^2} , \quad (1.20)$$

which, in Chap. 1, was given as a member of a sequence defining  $\delta(x)$ . Thus, initial conditions corresponding to a  $\delta(x)$  localization are included in (6.34a). One simply lets  $a \rightarrow 0$ . But then there is a price to pay, for by (6.37) we shall have  $\Delta p \rightarrow \infty$ . How about choosing some function other than a Gaussian? We have already seen in Example 1.1 that it is a property of the Fourier representation that for the product of the spread in wave vector space  $\Delta k$ , and in real space  $\Delta x$

$$\Delta k \Delta x \simeq 1 . \quad (6.40a)$$

Further, it can be shown that of all functions it is the Gaussian which yields the smallest value for this product, to wit:

$$\Delta k \Delta x = 1/2 . \quad (6.40b)$$

Obviously, these facts, inherent in wave mechanics, have broad implications. However, it will be advantageous to defer going further into the latter until we have solved our problem, i.e., found  $\psi(x, t)$  for  $t > 0$ .

Going back to (6.12), Schroedinger's time-independent equation for a free particle, one sees that

$$u_k(x) = e^{ikx}/\sqrt{2\pi} , \quad \infty > k > -\infty \quad (6.18)$$

give a complete set of free particle solutions when there is no need for any special boundary condition. The eigenfunctions of the time-independent total energy operator  $\mathbf{H}$  coincide with the eigenfunctions of the momentum operator  $\bar{\mathbf{p}}$ , see (6.17). Finding exactly when two or more operators turn out to have common eigenfunctions is really beyond the scope of our treatment; however, the interested reader will find that one of the problems does address this question.

As in (6.29), the time-dependent solution of Schroedinger's equation corresponding to  $u_k(x)$  is  $u_k(x)e^{-iE_k t/\hbar}$ . Or, following (6.25b), the most general solution for  $\psi(x, t)$  will be a linear combination of all possible solutions. Thus, for a free particle

$$\psi(x, t) = \int_{-\infty}^{\infty} A(k) u_k(x) e^{-iE_k t/\hbar} dk . \quad (6.41)$$

Just as for normal modes in classical physics, the coefficients  $A(k)$  are amplitudes to be determined from the initial conditions. In fact, looking at the derivation of the result for  $A(k)$  in (6.35), we have precisely found the coefficients for  $\psi(x, 0)$ . Therefore, there only remains an integration over  $k$  to give us  $\psi(x, t)$ . Substituting (6.35) for  $A(k)$  and (6.18) for  $u_k(x)$  into the general expression (6.41) we get

$$\psi(x, t) = \left(1/\sqrt{2\pi}\right) \left(2a^2/\pi\right)^{1/4} \int e^{ikx} e^{-i\omega_k t} e^{-(k-k_0)^2 a^2} dk . \quad (6.42a)$$

Using (6.15a), the formula for the angular frequency  $\omega_k = E_k/\hbar$  is given by

$$\omega(k) = \hbar k^2/2m . \quad (6.42b)$$

In our study of classical mechanics we were used to having a linear relationship between the wave number  $k$  and angular frequency  $\omega$ , for example, with waves on a stretched string. However, we have also encountered dispersive situations; the *phase speed*  $\omega/k$  of a light wave in a medium in Chap. 5 did depend on  $k$ , giving, for instance, prisms "dispersive power". In wave mechanics dispersion is evidently very marked.

The integral in (6.42a) is carried out by "completing the square". To reduce the algebra we introduce a complex quantity for the coefficient of  $k^2$  in the exponent, and also a "time"  $T$ , where

$$\Lambda^2 = a^2 + iT \quad (6.42c)$$

$$T = \hbar t/2m . \quad (6.42d)$$

With these symbols  $\psi(x, t)$  in (6.42a) becomes

$$\psi(x, t) = \left(1/\sqrt{2\pi}\right) (2a^2/\pi)^{1/4} e^{k_0^2 a^4/\Lambda^2} e^{-k_0^2 a^2} I$$

or

$$\psi(x, t) = (a^2/2\pi^3)^{1/4} e^{-ik_0^2 a^2 T/\Lambda^2} I, \quad (6.42e)$$

$$I = \int_{-\infty}^{\infty} e^{-(k\Lambda - k_0 a^2/\Lambda)^2 + ikx} dk. \quad (6.42f)$$

Here the factors  $\exp(k_0^2 a^4/\Lambda^2)$  and  $\exp -(k_0^2 a^4/\Lambda^2)$  have been introduced into (6.42e,f), respectively. Interchanging  $x$  and  $k$  and substituting  $\Lambda$  for  $a$ , in (1.18), the Fourier transform of a Gaussian, we get:

$$\int_{-\infty}^{\infty} e^{ikx} e^{-k^2 \Lambda^2} dk = (\sqrt{\pi}/\Lambda) e^{-x^2/4\Lambda^2}. \quad (6.42g)$$

Here it is important to notice that the real part of  $\Lambda^2, a^2$ , is always positive. Under these circumstances (1.18) may be extended, as is indicated in Chap. 1, to complex values of the parameter  $\Lambda$ . Further, in this case we may disregard a finite shift  $k_0 a^2/\Lambda^2$  in  $k$  when the range of integration in  $k$  is from plus to minus infinity. Thus, making the change from  $k$  to  $(k - k_0 a^2/\Lambda^2)$  in (6.42f), we have from (6.42g)

$$I = (\sqrt{\pi}/\Lambda) e^{-x^2/4\Lambda^2} e^{i(k_0 a^2/\Lambda^2)x}. \quad (6.42h)$$

Equations (6.42e,h) give us the sought after analytic solution for the wave function of a free particle  $\psi(x, t)$ .

We can, before dealing with the full complexity of the expressions for  $\psi(x, t)$ , discover some of the information we shall be wanting. Turning things around, we again note that for a free particle the eigenfunctions of the momentum operator  $u_k(x)$  are also eigenfunctions of the energy operator  $\mathbf{H}$ , giving us the expression  $A_k e^{-i(E_k/\hbar)t}$  for the time-dependent coefficients  $B_k(t)$  in the expansion (6.25b) for  $\psi(x, t)$  in momentum eigenfunctions:

$$\psi(x, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} B_k(t) u_k(x) dk. \quad (6.25b)$$

Thus, with  $|B_k(t)|^2$  independent of time,  $\bar{p}(t)$  and  $\Delta p(t)$  are, likewise, time-independent. We therefore have

$$\bar{p}(t) = \bar{p}(0) = \hbar k_0 \quad (6.43)$$

$$\Delta p(t) = \Delta p(0) = \pi/2a. \quad (6.44)$$

On the basis of classical physics this result is hardly surprising for a free particle, a point to which we shall return.

On the other hand  $\bar{x}(t)$  and  $\Delta x(t)$  are strongly time-dependent. To calculate this dependence we do not need the full  $\psi(x, t)$  as we do, in principle, to find the details of the momentum distribution. It suffices to know  $\psi^*(x, t)\psi(x, t)$ , the probability distribution of the particle. Looking at (6.42d) and (6.42g), we shall need:

$$\frac{1}{2} \left( \frac{1}{\Lambda^2} + \frac{1}{(\Lambda^2)^*} \right) = \frac{a^2}{\Delta} \quad (6.45a)$$

$$\frac{1}{2} \left( \frac{1}{\Lambda^2} - \frac{1}{(\Lambda^2)^*} \right) = \frac{-iT}{\Delta} \quad (6.45b)$$

$$(\Lambda\Lambda^*)^{-1} = (\Delta)^{-1/2} \quad (6.45c)$$

where

$$\Delta \equiv a^4 + T^2. \quad (6.45d)$$

With these results we find a perfect square in the exponent, that is, we find

$$\psi^*(x, t)\psi(x, t) = \left( a/\sqrt{2\pi\Delta} \right) e^{-(a^2/2\Delta)(x-\hbar k_0 t/m)^2}. \quad (6.45e)$$

Equation (6.45e) is our final result for the probability distribution of the free particle at time  $t$ . We can easily check that the distribution has remained normalized in time (6.23b), a consequence of the *unitary* property of the Schroedinger equation. The entire calculation for  $\psi^*\psi$  can be easily generalized to three dimensions, as can all of our subsequent discussion.

We can now find both  $\bar{x}(t)$  and  $\Delta x(t)$  rather directly from (6.45e). Proceeding as we did to find  $\bar{p}(0)$  in (6.36b), we see at once that

$$\begin{aligned} \bar{x}(t) &= \int_{-\infty}^{\infty} \psi^*(x, t)\psi(x, t)x dx \\ &= (\hbar k_0/m)t. \end{aligned} \quad (6.46a)$$

Also, remembering that  $\Delta x(0)$  in (6.38), associated with a normalized distribution going as  $e^{-x^2/2a^2}$ , was  $a$ , we have here, more generally, for the spread  $\Delta x(t)$ :

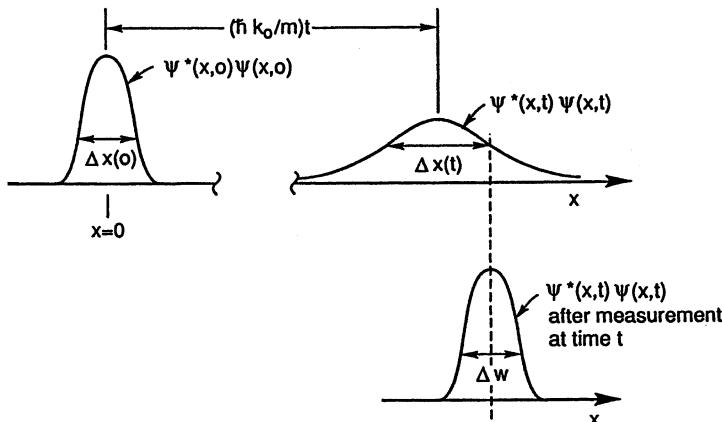
$$\Delta x(t) = \sqrt{\Delta}/a \quad (6.46b)$$

$$= \sqrt{a^2 + (\Delta p/m)^2 t^2}, \quad (6.46c)$$

where we have used (6.45d) for  $\Delta$ , and (6.42d) for  $T$ , and (6.44), which relates  $a$  to  $\Delta p$ .

The results of our calculation are summarized below comparing initial and later values, and in Fig. 6.2.

$t = 0$	$t > 0$
$\bar{x} = 0$	$\bar{x} = (\bar{p}/m)t$
$\Delta x = a$	$\Delta x = \sqrt{a^2 + (\Delta p/m)^2 t^2}$
$\bar{p} = \hbar k_0$	$\bar{p} = \hbar k_0$
$\Delta p = \hbar/2a$	$\Delta p = \hbar/2a$
$\Delta x \Delta p = \hbar/2$	$\Delta x \Delta p > \hbar/2$



**Fig. 6.2.** The probability distribution  $\psi^*\psi$  at  $t = 0$ , and at  $t > 0$ . The larger spread at finite  $t$  displays the statistical nature of wave mechanics. Also shown is  $\psi^*\psi$  after a measurement at time  $t$ . The wave function has “collapsed”

At first inspection what has happened is simple enough. The particle started at  $x = 0$ . We described it, the only way one can in wave mechanics, as a linear superposition of waves of varying wave vector  $k$ . The Dirac delta function  $\delta(x)$ , and the plane wave  $u_k(x)$  are the two extremes of possible descriptions being, as pointed out in Chap. 1, exact opposites.  $\delta(x)$  represents total localization, but the associated  $\Delta p$  goes to  $\infty$ , since all Fourier components are equally represented in the delta function.  $u_k(x)$ , a single harmonic wave, represents total delocalization,  $\Delta x \rightarrow \infty$ , with, however, a well defined momentum, i.e.,  $\Delta p \rightarrow 0$ . More generally, we have used a superposition of

waves called a *wave packet*. With the wave packet we found that no matter how we adjust parameters, we get a description in which

$$\Delta x \Delta p \geq \hbar/2. \quad (6.47)$$

By now such limitations are a familiar facet of wave formalisms, having their origin directly in Fourier analysis. We came across a similar limitation in Chap. 5:

$$\Delta E \Delta t \geq \hbar/2. \quad (5.5)$$

The subsequent motion of the free particle is entirely consistent with classical mechanics. The average momentum of the particle,  $\bar{p} = \hbar k_0$ , remains constant; the momentum distribution in  $k$  space does not change in time, and remains centered about  $k_0$ . This is what we expect for a *free* particle, i.e., when the potential  $V(x)$  does not have a gradient, when there are no forces.

In time  $t$  the peak of the probability distribution advances a distance  $(\bar{p}/m)t$ . The particle satisfies

$$\bar{x} = vt \quad (6.48)$$

where  $v$  is given by  $\hbar k_0/m$ , or, equivalently by  $\bar{p}/m$ .

However,  $\Delta x$  given by  $\sqrt{a^2 + (\Delta p/m)^2 t^2}$ , does change in time. The initial spread in momentum manifests itself as a spread in position at a later time.

The *Heisenberg uncertainty principle* states that the position of a particle and its momentum can not be measured simultaneously to an accuracy greater than that given in (6.47). Similarly the energy of a particle and the time during which the particle has that energy can not be ascertained to an accuracy exceeding that given in (5.5). We have already seen that (6.47) is the direct outcome of the original de Broglie relation (6.1) and Fourier analysis. What the Heisenberg principle states is that at the very precise level of measurement where (6.47) is relevant, the means of measurement will always affect the system to such an extent that accuracy exceeding  $\Delta x \Delta p = \hbar/2$  is unobtainable. We illustrate in Example 6.3 how the measurement of position with light (on the quantum level, by bouncing photons off the particle under scrutiny) disturbs the particle. When quantum physics was first formulated, physicists thought of many different hypothetical experiments to test the Heisenberg principle, but no experiment has ever been found to discredit the principle.

The limitation on measurement formulated by Heisenberg is a necessary condition for the ultimate validity of wave mechanics. Were it possible to make measurements which give data in excess of the limitations of (5.5) and (6.47), one could study such information and the subsequent evolution. Such knowledge would lie outside the scope of wave mechanics, which can only describe phenomena within the limitations. This would not, in itself, invalidate the theory, but it would indicate the existence of a yet more basic formulation, and some further way of understanding nature. Rather, we see that quantum mechanics has an internal consistency; in fact, it has correctly

predicted such an incredible wealth of physical phenomena on the atomic and sub-atomic level, where Newtonian mechanics has been found to break down, to have become on this level the accepted replacement.

If we now focus on the free particle at time  $t$ , where  $\Delta x \Delta p_x$  is definitely greater than  $\hbar/2$ , here measurements will allow more precise simultaneous assessment of the dynamic variables than our calculation was able to predict, i.e., to within  $\hbar/2$ . Thus, there will be more than one possible outcome from the measurement,  $\psi^*(x, t)\psi(x, t)$  representing a statistical distribution. We can, for instance, take a Geiger counter (a device that counts particles individually) with a window opening  $\Delta w$  set at considerably less than  $\Delta x(t)$ , and scan the probability distribution at time  $t$ . We will then find that, on the average, repeating the experiment many times, the total number of counts (per particle) will be distributed according to the function  $\psi^*(x, t)\psi(x, t)$ ;  $\psi^*(x, t)\psi(x, t)\Delta w$  is the likelihood that the particle will be found at  $x$  within a window opening of width  $\Delta w$ .

After the measurements, the indeterminacy of the wave function may be back to what it was at  $t = 0$ ; we simply set  $\Delta w = \Delta x(0)$ . This is sometimes called a “collapse” of the wave function (Fig. 6.2), an alteration to which we shall refer again.

Lastly, we note that our calculation of the propagation of a free particle, however strong the relation of the results to classical particle mechanics, did, in fact, proceed by the way of *wave* mechanics; the propagation of the particle was represented by the motion of a wave packet. This *wave particle duality* can be illustrated most vividly by a rather different physical situation described in the next section.

## 6.4 Wave-Particle Duality and Loss of Determinism

The two-slit diffraction experiment of the previous chapter can be repeated with particles. A beam of particles, say, electrons, is passed through two horizontal slits in a baffle. The vertical distribution of arriving electrons can be mapped out with the use of a Geiger counter, just as the distribution of free particles could be ascertained in the previous section. In its distribution, the intensity of the diffracted electron beam will exactly resemble the pattern obtained in Fig. 5.12 for diffracted light. Now, the Geiger counter is a device that records the arrival of *particles*, it either clicks or it does not; there are no half-clicks. If a bunch of heavy particles, say, pellets, had been directed through the slits, the distribution along a vertical receiving plane could not possibly resemble the observed complex diffraction pattern. The pellets would simply be found to arrive along two well-defined, well-separated lines, as predicted by classical mechanics. The departure from such a distribution, namely, the complex diffraction pattern, may be taken as indicating that when an electron goes through one of the slits it knows that the other slit is there as well! – the wave-particle duality at its most evident.

At this point one becomes curious as to what would happen if the Geiger counter, with all its complexity, were formally included in the problem, i.e., the whole mechanism became part of the potential  $V$  in the Schroedinger wave equation. In that case we would do our physics in the usual deterministic fashion, simply solving a (nonsimple!) wave equation, just as we do in classical physics. The wave-particle duality, the indeterminism, and the “collapse” of the wave function would disappear from the calculation. Would they disappear altogether? The answer is they would be pushed to the next level of observation. In the end, to quote from Heisenberg’s *Physical Principles of the Quantum Theory* (see Further Reading for Chap. 6): “we should be forced, for example, to include our own eyes as part of the system, and so on. The chain of cause and effect could be quantitatively verified only if the whole universe were considered as a single system – but then physics has vanished, and only a mathematical scheme remains. The partition of the world into observing and observed system prevents a sharp formulation of the law of cause and effect. (The observing system need not always be a human being; it may also be an inanimate apparatus, such as a photographic plate.)”

In practice, in view of the difficulty of solving the Schroedinger equation, only very simple problems can be analyzed. The rest of the work of interpretation of experiments is as much an art as a science. Nevertheless, physicists have successfully managed to quantitatively understand a wealth of information on the quantum level of nature. We demonstrate this in the next section, the last section of this chapter.

Wave mechanics, when first discovered, caused consternation among many physicists. A wave function which, on the one hand, contained information of only statistical value had, on the other, to be carried through calculations in all its entirety (as in a multiple slit experiment), yet collapsed with each measurement. Schroedinger himself posed the famous cat paradox (Schroedinger’s Cat). A cat is shut up in a box together with a small amount of radioactive substance such that in an hour there is a 50% chance of a radioactive decay. If there is a decay, an apparatus is triggered which releases a deadly gas. The  $\psi$  function of the system at the end of an hour would express this situation by being a linear combination with equal coefficients of components describing a living and a nonliving animal. Is the cat really neither alive nor dead until an observer has looked into the box? A different but related question is, will there be any evidence upon opening of the box of strange behavior attributable to the mixed wave function, as there is with the composite wave function of an electron that has been diffracted by two slits?

Today most physicists are no longer troubled by such questions. In particular, an animal or human brain (or even one transistor), is complex; so many neural processes go into any human decision, that as regards individual life, or recorded history, any effects of quantum fluctuations are almost certainly completely obliterated. As regards Schroedinger’s cat, the wave function would contain a tremendous number of eventualities, almost all of them leading to one or the other of the large scale outcomes, with only a totally negligible number of eventualities leading to something strange.

Let us consider a simpler example. Suppose we measure simultaneously the position and momentum of the moon, with the inevitable Heisenberg uncertainty. It is posed as one of the Problems to determine how long it will take for the moon to come off-course by one kilometer. In enough time (eons), the initial quantum

uncertainty can lead to such an effect. However, if it does, it will not come as a surprise, since, inevitably, we will take a look now and then. As in any statistical process, monitoring will limit the statistical spread in the predictions. At the same time there is nothing to support the notion that the collapsing of the wave functions in themselves affect future outcomes. It is the measurements leading to the collapse that have effects (Example 6.3).

In summary, when dealing with the constituents of nature, very light particles or photons, it is necessary to relax certain concepts based on direct experience. Here visualization of particles as concrete objects, so useful in Newtonian mechanics, must be modified by notions of wave particle duality. The German physicist Max von Laue (1879–1960) suggested that a quantum mechanical object is to be viewed as a wave when looking towards the future, but as a particle when considering the past! Alternately, in answer to *where*, we have a wave, in answer to *how*, we have a particle. However, despite the unpredictability which appears in quantum mechanics when the world is split into a subject and object, we can integrate the quantum results into our picture of an orderly universe. Most “hardware” is sufficiently massive or complex (redundant) for quantum fluctuations to have any tangible effect. In his famous “correspondence principle” Niels Bohr emphasized the requirement that when physical situations approach a nonquantum limit, usually because there are present many quanta of energy or angular momentum, quantum theory results must go into those of classical physics. This does happen in quantum theory as it does in experiment.

## 6.5 Driven Motion and the Green's Function for the Harmonic Oscillator

We have already seen that, strictly, the Schrödinger equation applies to the universe as a whole. The choice of a workable model, for example, doing calculations ignoring the apparatus of observation requires expertise in interpreting the results. Nevertheless, the concept of an external driving agent, providing an intuitive dissection of problems in classical physics is also useful in quantum mechanics. For example, if one were studying infrared absorption by the elastic waves of an ionic crystal, one would assume a given, fixed amplitude for the applied field. One could safely neglect that oscillations of the crystal ions, driven by the field were producing additional radiation, which might even be reacting back on the original source. This assumption is in contrast to what happens in, say, a laser, where the induced radiation is essential to the laser's operation. In addition, for the laser one needs to quantize the electromagnetic field, introducing photons (see next section). However, in calculating the effect of the infrared on the lattice waves one may work with an external driving field, and treat that field as in classical physics. The concept of driven as contrasted to natural motion is, thereby, carried over from classical to quantum mechanics.

The dynamics of all the phenomena studied in the previous chapters on classical mechanics rested on the time dependent behavior of the harmonic oscillator. Here we will develop the necessary aspects of the quantum physics of harmonic oscillators, and, subsequently, the Green's function technique for treating the driven motion. This analysis will contribute to our goal of integrating quantum mechanics with the previous exposition of classical wave physics.

Possibly the system with the simplest time dependent behavior in quantum mechanics is one with only two energy levels. For example, problems involving electron "spin" (intrinsic angular momentum) may require just two states: the magnetic moment of the electron can point either parallel or anti-parallel to an applied magnetic field. In practice these two states do not form a complete set for the electron, but the other eigenstates are so far away in energy compared to the spacing between the two states, that the exclusion of the other states will have negligible influence on the results of the calculation. The two level problems can be solved quite directly by expanding the time dependent wave function in energy eigenstates, letting the time be carried by coefficients  $c_n(t)$ ,  $n = 1, 2$ . These simplest systems can be of great importance technically. Example (6.4) explores the basis of the nuclear magnetic resonance (NMR) imaging technique, so useful in medicine. Problem (6.19) deals with the origin of the resonance response of the electronic polarizability of dielectrics.

Let us return to the study of the driven harmonic oscillator, the subject of this section. The total energy operator  $\mathbf{H}_0$  for a free harmonic oscillator is given by:

$$\mathbf{H}_0 = \mathbf{p}^2/2m + 1/2kx^2 \quad (6.49)$$

Here the first term on the right is the kinetic energy, the second term the potential energy stored in the spring. In quantum mechanics,  $p$ , the momentum of the mass  $m$  becomes the operator  $(\hbar/i)/\partial x$  with one dimension  $x$ .  $k$  is the spring constant.

The harmonic oscillator can be treated by means of two very significant operators, defined as:

$$\mathbf{a} \equiv i/(2m\hbar\omega)^{1/2}(\mathbf{p} - im\omega x) \quad (6.50a)$$

$$\mathbf{a}^\dagger \equiv -i/(2m\hbar\omega)^{1/2}(\mathbf{p} + im\omega x) \quad (6.50b)$$

Here the superscript (dagger) of the operator in (6.50b) denotes that this operator is the *adjoint* of the original operator defined in (6.50a). For any operator  $\mathbf{Q}$  one can define an operator  $\mathbf{Q}^\dagger$  by the equation:

$$\int_{-\infty}^{\infty} v^*(x)\mathbf{Q}^\dagger w(x)dx = \int_{-\infty}^{\infty} [\mathbf{Q}v(x)]^* w(x)dx \quad (6.51)$$

where  $v, w$  are any good functions of  $x$ , and the asterisks (\*) denote complex conjugates, as usual. By doing "integrations by parts" when treating the momentum operator  $\mathbf{p}$  the reader can easily show that  $\mathbf{a}$  and  $\mathbf{a}^\dagger$  are, indeed,

adjoints of each other. While, possibly, unfamiliar to the reader, (6.51) is, in fact, an integral part of the mathematics of quantum mechanics. An operator which has  $\mathbf{Q}^+ = \mathbf{Q}$  is said to be Hermetian. As already stated in Sect. 6.2, Hermetian operators have real eigenvalues, see Problem (6.15).  $\mathbf{a}$  and  $\mathbf{a}^+$  are decidedly not Hermetian.

If we solve (6.50a,b) for  $\mathbf{p}$  and  $x$ , we obtain:

$$\mathbf{p} = \left( \frac{m\hbar\omega}{2} \right)^{1/2} i(\mathbf{a}^+ - \mathbf{a}) \quad (6.52a)$$

$$\mathbf{x} = \left( \frac{\hbar}{2m\omega} \right)^{1/2} i(\mathbf{a}^+ + \mathbf{a}) \quad (6.52b)$$

Substituting these expressions for  $\mathbf{p}$  and  $x$  back into (6.49) for  $\mathbf{H}_0$  we obtain after some algebra:

$$\mathbf{H}_0 = \left( \frac{1}{2} \right) \hbar\omega (\mathbf{a}^+ \mathbf{a} + \mathbf{a} \mathbf{a}^+) \quad (6.53)$$

We note that  $\mathbf{a}, \mathbf{a}^+$  being operators, their order of operating is important, that is,  $\mathbf{a}$  and  $\mathbf{a}^+$  do not commute (see Problem 6.7). In fact, from their definitions (6.50a,b) one finds the *commutation relation*:

$$\mathbf{a}\mathbf{a}^+ - \mathbf{a}^+\mathbf{a} = 1 \quad (6.54)$$

Accordingly, we write (6.53) as

$$\mathbf{H}_0 = \hbar\omega (\mathbf{a}^+ \mathbf{a} + \frac{1}{2}) \quad (6.55)$$

As it happens, one can easily find the energy eigenvalues of the Hamiltonian (6.49), or its equivalent (6.55) without actually solving Schroedinger's differential equation. Suppose that one has available an energy eigenstate  $u_n(x)$ , where, anticipating,  $n$  is 0 or an integer:

$$\mathbf{H}_0 u_n(x) = E_n u_n(x) \quad (6.56)$$

$E_n$  is the corresponding energy eigenvalue. Now, consider the function  $\mathbf{a} u_n(x)$ . One has from (6.55):

$$\mathbf{H}_0 [\mathbf{a}^+ u_n(x)] = \hbar\omega [\mathbf{a}^+ \mathbf{a} \mathbf{a}^+ + \mathbf{a}^+ / 2] u_n(x) \quad (6.57a)$$

$$= \hbar\omega (\mathbf{a}^+ \mathbf{a}^+ \mathbf{a} + \mathbf{a}^+ / 2) u_n(x) \quad (6.57b)$$

$$= \mathbf{a}^+ \hbar\omega (\mathbf{a}^+ \mathbf{a} + \frac{1}{2}) u_n(x) + \hbar\omega \mathbf{a}^+ u_n(x) \quad (6.57c)$$

$$= (E_n + \hbar\omega) \mathbf{a}^+ u_n(x) \quad (6.57d)$$

where one has used (6.54) to obtain (6.57b), and (6.56) to go from (6.57c) to (6.57d). (6.57d) shows that if  $u_n(x)$  is an energy eigenfunction with eigenvalue  $E_n$ ,  $\mathbf{a}^+ u_n(x)$  is likewise an eigenfunction with eigenvalue  $(E_n + \hbar\omega)$ . If one proceeds similarly, focusing on  $\mathbf{a} u_n(x)$ , one learns that  $\mathbf{a} u_n(x)$  is also an eigenfunction, with eigenvalue  $(E_n - \hbar\omega)$ .

Physically, one must suppose that a ground state, call it  $u_0(x)$ , exists. The classical physics equivalent of the ground state is when the oscillator is

standing still at its equilibrium point, and has then zero energy. However, some small energy,  $E_0$ , must be associated with  $u_0$  to allow the uncertainty principle  $\Delta x \Delta p \geq \hbar/2$  to be satisfied. We have just learned that the operator  $a$ , when applied to an eigenstate leads to a state with yet lower energy. To terminate this process one will have to have:

$$au_0(x) = 0 \quad (6.58)$$

(If a differential equation method is used, solution likewise involves a termination condition.) Moreover, from what we have learned about the eigenvalues of the oscillator we have proportions:

$$au_n(x) \propto u_{n-1}(x) \quad n = 1, 2, 3, \dots \quad (6.59a)$$

$$a^+ u_n(x) \propto u_{n+1}(x) \quad n = 0, 1, 2, 3, \dots \quad (6.59b)$$

and the energy eigenvalues will be given by:

$$E_n = E_0 + n\hbar\omega \quad n = 0, 1, 2, 3, \dots \quad (6.60)$$

The original equations (6.54) and (6.55) along with (6.60) for the eigenvalues can now be solved uniquely by letting:

$$a^+ u_n(x) = (n + 1)^{1/2} u_{n+1}(x) \quad n = 0, 1, 2, 3, \dots \quad (6.61a)$$

$$au_n(x) = n^{1/2} u_{n-1}(x) \quad n = 0, 1, 2, 3, \dots \quad (6.61b)$$

where the operators  $a^+, a$  are now seen to be raising and lowering operators. Finally, substituting (6.61a,b) into the equation (6.55) for the Hamiltonian of the oscillator yields for the eigenvalue equation (6.56):

$$H_0 u_n(x) = \hbar\omega (n + \frac{1}{2}) \quad n = 0, 1, 2, 3, \dots \quad (6.62)$$

(6.61a,b), with  $a$  and  $a^+$  defined in (6.50a,b) give full information equivalent to knowing the eigenstates  $u_n(x)$ , as we shall see.

It may come as a surprise that one is able to solve the problem this way, without explicitly solving Schroedinger's time independent equation. The fact is that in quantum mechanics systems are fully determined once the Hamiltonian (energy operator) is given, and the component operators are defined by their commutation relations, here (6.54). The skeptical reader is, however, referred to Problem 6.16. Here the ground state solution  $u_0(x)$  of Schroedinger's equation is explicitly given, the relation  $au_0(x) = 0$  demonstrated, and the ground state energy,  $\hbar\omega/2$  shown to be compatible with the requirements of the uncertainty principle.

Having solved the eigenvalue (normal mode) equation (6.56), we can solve fully the initial value problem for a harmonic oscillator undergoing natural motion. We revert to an expansion like that given in (6.25a) along with (6.26) which, as we have seen, solve the Schroedinger time dependent equation. However, for the more general case of driven motion calculations become more difficult. There is, however, another way of solving time dependent problems in quantum mechanics, which we can apply. This method will prove to be particularly helpful in displaying the relation between the classical and quantum

mechanical way of treating driven motion. We have already learned that the raising and stepping down operators  $a^+$ ,  $a$  are extremely useful. Can they, perhaps, be generalized to carry the time dependence? In fact, time dependence in quantum mechanics can be treated in two equivalent ways, or according to two different *pictures*. Up to now we have used the Schroedinger picture. Here the time dependence is contained in the wave function  $\Psi(x, t)$ . We can write the time dependent Schroedinger equation (6.4) for  $\Psi(x, t)$  in the form:

$$i\hbar \frac{\partial \Psi(x, t)}{\partial t} = \mathbf{H}(x, t)\Psi(x, t) \quad (6.63a)$$

$$\mathbf{H}(x, t) = \mathbf{H}_0(x) + V(x, t) \quad (6.63b)$$

Since, we are interested in external driving forces that are time dependent, we must allow for the possibility that  $V(x, t)$ , and, hence, also  $\mathbf{H}(x, t)$  are time dependent. As explained in Sect. 6.1, the Schroedinger equation holds for such an extension of  $V$ .

In the Heisenberg picture, the time dependence is transferred from the wave functions to the quantum mechanical operators. This is done by means of an operator  $\mathbf{U}(t)$  which is defined as taking the wave function from the initial time, say,  $t = 0$ , to the desired final time  $t$ :

$$\boxed{\Psi(x, t) \equiv \mathbf{U}(x, t)\Psi(x, 0)} \quad (6.64)$$

$\mathbf{U}(t)$  is to be found from Schroedinger's time dependent wave equation (6.63a).

In solving (6.63a) we must be careful to remember that  $\mathbf{H}$  is an operator. Going back to the definition of a derivative we devide the time interval 0 to  $t$  into  $N$  equal intervals,  $\Delta t$ , ( $\Delta t = t/N$ ), and ultimately let  $N \rightarrow \infty$ . For the first interval we can now rewrite (6.63a):

$$\Psi(x, \Delta t) - \Psi(x, 0) \simeq (-i/\hbar)\mathbf{H}(0)\Delta t\Psi(x, 0) \quad (6.65a)$$

yielding from (6.63), the definition of  $\mathbf{U}(t)$ :

$$\mathbf{U}(\Delta t) \simeq 1 - (i/\hbar)\mathbf{H}(0)\Delta t \quad (6.65b)$$

$$\simeq e^{(i/\hbar)\mathbf{H}(0)\Delta t} \quad (6.65c)$$

$$\equiv 1 - (i/\hbar)\mathbf{H}(0)\Delta t + (i/\hbar)^2\mathbf{H}(0)\Delta t^2 \dots \quad (6.65d)$$

with, clearly,  $\mathbf{U}(0) = 1$ . As long as  $\Delta t$  is small enough there is no distinction between the two forms (6.65b), (6.65c). However, to obtain solutions for finite time intervals we shall need, as  $\Delta t$  gets smaller and smaller, to take products of an infinite number of factors such as those that appear on the right of (6.65b,c). We can only take such products and come out with the correct limit by using the second form, (6.65c). Proceeding with our calculation, increasing the time step by step, we obtain:

$$\mathbf{U}(t) = \exp [-(i/\hbar)\Delta t \mathbf{H}(\overline{N-1}\Delta t)] \dots \exp [-(i/\hbar)\Delta t \mathbf{H}(0)] \quad (6.66)$$

(6.66) tells us immediately that:

$$\mathbf{U}(t) \equiv \mathbf{U}(t, 0) = \mathbf{U}(t, t')\mathbf{U}(t', 0) \quad t \geq t' \geq 0 \quad (6.67)$$

One result useful in handling exponentials with operators, such as (6.65b) is:

$$\exp \mathbf{A} \exp \mathbf{B} = \exp(\mathbf{A} + \mathbf{B} + \frac{1}{2}[\mathbf{A}, \mathbf{B}]) \quad (6.68)$$

where  $\mathbf{A}$  and  $\mathbf{B}$  are operators, and  $[\mathbf{A}, \mathbf{B}]$ , their commutator, is a scalar (ordinary number, or vanishes). For a proof the reader is referred to the book by Messiah (Sect. 12.2.1) listed at the end of this chapter. If now  $\mathbf{H}$  is independent of  $t$ , for example, if it equals  $\mathbf{H}_0$ , defined in (6.49), then, obviously, we obtain:

$$\mathbf{U}(t) = e^{-(i/\hbar)\mathbf{H}_0 t} \quad (6.69)$$

If  $\mathbf{H}$  does depend on  $t$ , special circumspection is called for when applying (6.66). Here we will not be called upon to perform explicit calculations when this last is the case.

The transforming operator,  $\mathbf{U}$  has important properties, which we would like now to establish. First:

$$\mathbf{U}^+(t) = \mathbf{U}^*(t) \quad (6.70a)$$

where  $\mathbf{U}^+$  is the adjoint operator of  $\mathbf{U}$ , see (6.51) for the definition of "adjoint", whereas  $\mathbf{U}^*$  is the complex conjugate of  $\mathbf{U}$ . Noting that the potential in (6.63) is a real scalar, we have

$$\int_{-\infty}^{\infty} [\mathbf{H}v(x)]^* w(x) dx = \int_{-\infty}^{\infty} v^*(x) \mathbf{H}w(x) dx \quad (6.71)$$

Here we have applied integration by parts (chain rule) twice to deal with the kinetic energy. The integrals vanish at  $\pm\infty$  owing to wave functions  $v, w$  being good functions. From (6.71),  $\mathbf{H}$  is a Hermitian operator, as was to be expected since it has real eigenvalues. One can now see that any polynomial of  $i\mathbf{H}$  satisfies a relationship like (6.70a). Since, we can always expand  $\mathbf{U}(t)$ , as given in (6.66), into such polynomials, (6.70a) is going to hold.

The operator  $\mathbf{U}$  will also possess an inverse operator,  $\mathbf{U}^{-1}$ , defined by the equation

$$\mathbf{U}(t)\mathbf{U}^{-1}(t) \equiv 1 \quad (6.70b)$$

Inspection of (6.66), which gives  $\mathbf{U}(t)$  will reveal that:

$$\mathbf{U}^{-1}(t) = \mathbf{U}^*(t) \quad (6.70c)$$

As a result we finally have:

$$\mathbf{U}^+(t) = \mathbf{U}^{-1}(t) \quad (6.70d)$$

(6.70d) tells us that  $\mathbf{U}(t)$  is a unitary matrix because its adjoint is equal to its inverse. Alltogether we have four relationships for the operator  $\mathbf{U}$ : (6.70a) relating  $\mathbf{U}$  to its complex conjugate  $\mathbf{U}^*$ , ii) (6.70b) defining the inverse

operator, iii) (6.70c), and iv) (6.70d) telling us that  $\mathbf{U}$  is unitary. Lastly, for completeness we note that

$$\mathbf{U}^+(t) \neq \mathbf{U}(t) \quad (6.70e)$$

$\mathbf{U}$  is not a Hermitian operator.

We are now in a position to carry out our aim of converting from the Schroedinger to the Heisenberg picture. Explicitly or implicitly quantum mechanical calculations in the Schroedinger picture involve taking overlap integrals  $I$ , where

$$I \equiv \int_{-\infty}^{\infty} \phi^*(x, t) \mathbf{Q} \Psi(x, t) dx \quad (6.72)$$

$\phi$  and  $\Psi$  are any two functions satisfying the Schroedinger time dependent equation, and  $\mathbf{Q}$  is a quantum mechanical operator. With the help of (6.64), defining the operator  $\mathbf{U}(t)$ , the integral  $I$  becomes:

$$I = \int_{-\infty}^{\infty} [\mathbf{U}(t)\phi(x, 0)]^* \mathbf{Q} \mathbf{U}(t) \Psi(x, 0) dx \quad (6.73a)$$

$$= \int_{-\infty}^{\infty} \phi^*(x, 0) \mathbf{U}^{-1}(t) \mathbf{Q} \mathbf{U}(t) \Psi(x, 0) dx \quad (6.73b)$$

$$= \int_{-\infty}^{\infty} \phi^*(x, 0) \mathbf{Q}(t) \Psi(x, 0) dx \quad (6.73c)$$

with

$$\boxed{\mathbf{Q}(t) \equiv \mathbf{U}^{-1}(t) \mathbf{Q} \mathbf{U}(t)} \quad (6.73d)$$

Here (6.51), the definition of an adjunct operator has been combined with (6.70d) to go from (6.73a) to (6.73b). We have carried out our aim of transferring the time dependence from the Schroedinger wave function to Heisenberg operators  $\mathbf{Q}(t)$ , defined in (6.73d). The Schroedinger operator  $\mathbf{Q}(0)$ , i.e. the operator at the initial time  $t = 0$ , will continue to be designated as simply  $\mathbf{Q}$ .

Noting the form (6.66) of  $\mathbf{U}(t)$ , we have the relation:

$$\mathbf{U}^{-1}(t) \mathbf{H}(t) \mathbf{U}(t) = \mathbf{H}(t) \quad (6.74)$$

There is only one  $\mathbf{H}(t)$ ; it is the same in both pictures.

To calculate the time dependence of Heisenberg operators we find from (6.73d) and with the help of (6.70a) that:

$$\frac{\partial \mathbf{Q}(t)}{\partial t} = \frac{\partial \mathbf{U}^*(t)}{\partial t} \mathbf{Q} \mathbf{U}(t) + \mathbf{U}^*(t) \mathbf{Q} \frac{\partial \mathbf{U}(t)}{\partial t} \quad (6.75a)$$

if  $\mathbf{Q}$  does not depend explicitly on time. If we now combine Schroedinger's equation (6.63a) with (6.64), the definition of  $\mathbf{U}(t)$ , we get the operator equation:

$$i\hbar \frac{\partial \mathbf{U}(t)}{\partial t} = \mathbf{H}(t)\mathbf{U}(t) \quad (6.75b)$$

Combining (6.75a) with (6.75b) and the complex conjugate of equation (6.75b) we obtain:

$$\frac{\partial \mathbf{Q}(t)}{\partial t} = \frac{-1}{i\hbar} \{ \mathbf{H}(t)\mathbf{U}(t)\mathbf{Q}\mathbf{U}^{-1}(t) - \mathbf{U}^{-1}(t)\mathbf{Q}\mathbf{U}(t)\mathbf{U}^{-1}(t)\mathbf{H}(t)\mathbf{U}(t) \} \quad (6.75c)$$

or, with the help of (6.73d) and (6.74)

$$\frac{\partial \mathbf{Q}}{\partial t} = \frac{i}{\hbar} [\mathbf{H}(t), \mathbf{Q}(t)] \quad (6.75d)$$

In the second term of (6.75c) an extra factor  $\mathbf{U}(t)\mathbf{U}^{-1}(t)$  has been inserted, and (6.74) has been invoked in going to (6.75d). The square brackets in the final result (6.75d) denote the commutator.

We are ready to tackle the driven harmonic oscillator. We shall have:

$$V(x, t) = -xF(t) \quad (6.76a)$$

where  $F(t)$  is the applied time dependent force. It is being assumed that this force does not depend on  $x$ , the location of the oscillating mass. We can, thus, write for the Hamiltonian:

$$\mathbf{H}(t) = \mathbf{H}_0 + V(x, t) \quad (6.63b)$$

$$\mathbf{H}_0 = \mathbf{p}^2/2m + kx^2/2 \quad (6.49)$$

$$\mathbf{H}(t) = \frac{-\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + kx^2/2 - xF(t) \quad (6.76b)$$

Using the equation for the time rate of change of an operator (6.75d) we have:

$$\frac{\partial x}{\partial t} = \frac{i}{\hbar} [\mathbf{H}(t), x(t)] \quad (6.77a)$$

$$= i/(2m\hbar) [\mathbf{p}^2(t), x(t)] \quad (6.77b)$$

$$= \mathbf{p}(t)/m \quad (6.77c)$$

since  $[\mathbf{p}, x]$  does not depend on time. Furthermore,

$$\frac{\partial^2 x(t)}{\partial t^2} = \frac{1}{m} \frac{\partial \mathbf{p}(t)}{\partial t} \quad (6.77d)$$

and,

$$\frac{\partial \mathbf{p}(t)}{\partial t} = \frac{i}{\hbar} \left[ \frac{1}{2} kx^2(t) - x(t)F(t), \mathbf{p}(t) \right] \quad (6.77e)$$

again from (6.75d). We get:

$$\frac{\partial \mathbf{p}}{\partial t} = -kx(t) + F(t) \quad (6.77f)$$

Finally, combining (6.77d) with (6.77f) we have:

$$m \frac{\partial^2 x}{\partial t^2} + kx = F(t). \quad (6.77g)$$

Again, the equation of motion (6.77g) is identical with the classical equation (2.1), when damping is omitted in the latter. However, it must be remembered that here  $x(t)$  is a quantum mechanical operator.

Nevertheless, (6.77g) can be solved by the Green's function method. It is conventional to introduce a change of scale when doing the problem in quantum physics. The unit of length is:

$$x_0 \equiv \left( \frac{\hbar}{2m\omega} \right)^{1/2} \quad (6.78)$$

(6.77g) is rewritten as:

$$\frac{\partial^2}{\partial t^2} \left( \frac{x}{x_0} \right) + \omega^2 \left( \frac{x}{x_0} \right) = f(t)\omega \quad t > 0 \quad (6.79a)$$

We have rescaled  $F(t)$  as:

$$F(t) = mx_0\omega f(t) \quad (6.79b)$$

Now,  $x$  is an operator with expectation value

$$\bar{x}(t) = \int_{-\infty}^{\infty} \Psi^*(x, 0)x(t)\Psi(x, 0)dx \quad (6.80a)$$

Taking expectation values of both sides, one finds that (6.79a) becomes:

$$\frac{\partial^2}{\partial t^2} \left( \frac{\bar{x}}{x_0} \right) + \omega^2 \left( \frac{\bar{x}}{x_0} \right) = f(t)\omega \quad t > 0 \quad (6.80b)$$

In view of (6.80a,b), we define the Green's function  $G(t)$  as the solution to the equation:

$$\frac{d^2 G(t)}{dt^2} + \omega^2 G(t) = \omega\delta(t) \quad (6.81)$$

Here  $G(t)$  is dimensionless as a result of the rescaling of  $F(t)$ , it being  $f(t)$  which has been replaced by the delta function  $\delta(t)$ .

We recall from Sect. 2.5 that to obtain a solution for  $G$  we must specify the initial value and the initial derivative. Here we have again ( $t_0 = 0$ ) that

$$G(t) = 0 \quad t \leq 0 \quad (6.82a)$$

As in Chap. 2 we can integrate (6.81) over time from just below  $t = 0$ , ( $t = 0^-$ ) to just above  $t = 0$ , ( $t = 0^+$ ). We find:

$$\left( \frac{dG(t)}{dt} \Big|_{t=0^+} - \frac{dG(t)}{dt} \Big|_{t=0^-} \right) = \omega \quad (6.82b)$$

where the integral  $\omega^2 \int_{0^-}^{0^+} G(t)dt$  goes to zero, and (1.19b), the normalization of the delta function, was used to obtain the right side. Since, from (6.82a) the

second term on the left of (6.82b) vanishes, we have as the second condition on  $G(t)$  that

$$\left. \frac{dG(t)}{dt} \right|_{t=0^+} = \omega \quad (6.82c)$$

Concurrently, we see from (6.77c), after taking expectation values that:

$$\frac{d\bar{x}}{dt} = \frac{\bar{p}(t)}{m} \quad (6.83a)$$

or,

$$\frac{dG(t)}{dt} = \frac{\bar{p}(t)}{mx_0} \quad (6.83b)$$

From (6.82c) and (6.83b) we can write:

$$\bar{p}(0) = \omega mx_0 \quad (6.83c)$$

Thus, (6.83c) is equivalent physically to the initial condition on  $G(t)$  that the momentum absorbed on the average by the oscillator, i.e.  $\omega mx_0$  in (6.82c) is equal to the impulse imparted by the delta function force, see (6.79b).

As in Sect. 2.5 the solution

$$G(t) = 0 \quad t < 0 \quad (6.84a)$$

$$G(t) = \sin \omega t \quad t > 0 \quad (6.84b)$$

satisfies the initial conditions, here (6.82a) and (6.82c), as well as the equation of motion (6.81). Indeed, (6.84a,b) are the functions (2.30a,b) of  $t$ , rescaled.

We see from (2.32) that, with the damping missing, the Fourier transform  $G(\omega)$  of  $G(t)$  in (6.84a,b) is given by:

$$G(\omega) = -1/\sqrt{2\pi} \frac{\omega}{\omega^2 - \omega_0^2} \quad (6.85)$$

Damping is also found in a full quantum mechanical analysis, with  $G(t)$  ultimately going to zero as  $t$  gets large. To get damping we would, at least, in principle, have to include in the Hamiltonian operator the mechanism whereby the damping occurs, for example, weak coupling to other oscillators that model the environment. The analytic results, then, are similar but not identical with those contained in equation (2.32) for the classical Green's function  $G(\omega)$  of a damped oscillator.

Of interest here, as in Example 6.4 is that the resonance phenomenon, so striking in classical physics carries over into quantum mechanics. Indeed, given Bohr's correspondence principle of quantum mechanics going under appropriate conditions of many quanta into classical mechanics, this outcome is not unexpected. However, we must be very careful not to dismiss quantum mechanics as an academic exercise whose practical consequences

are anticipated by classical physics. We hope to make this clear in the next section, which will indicate how much more of nature's functioning we can understand as quantum physicists. The Green's function, as we have been learning about it will turn out to play an important part in the emerging development.

## 6.6 Scope of Quantum Mechanics

The unfolding of quantum mechanics has led to an understanding of an ever wider range of physical phenomena. In Sect. 6.1 we began to discuss how Schroedinger obtained the entire range of bound quantum states,  $1s$ ,  $2s$ ,  $2p$ ,  $3s$ ,  $3p$ ,  $3d\dots$  etc., of a hydrogen atom as the eigenstates (normal modes), of his time-independent equation. Since the structure of the entire periodic table of the elements is the direct extrapolation of these states, one begins to see that the structure of matter on the chemical level is to be understood as a consequence of wave mechanics. Besides the electronic structure of the atoms themselves, quantum mechanics also elucidates the nature of the chemical bonds in molecules and solids. Condensed matter physics goes on to describe the many physical properties of matter: crystal structure, electric conductivity in metals and semiconductors, magnetism, the superfluidity of liquid helium II and, still evolving, the mechanism of superconductivity, to name a few.

Returning to atomic physics, one finds that the energy levels in hydrogen are influenced not only by the details of the charge distribution  $\Psi^*\Psi$  as characterized by an electron "radius" (principle quantum number  $n$ ), but directly by the angular momentum of the electron (quantum number  $l$ ). Every electron possesses an *intrinsic magnetic moment* due to an intrinsic "spin", rotation about its own axis (see below). What happens is that, as predicted by special relativity, the electron moving in the Coulomb field of the nucleus also sees a magnetic field. This field wants to align the electron's intrinsic magnetic moment with it, leading to a shift in energy level dependent on the quantum number  $l$ . Such shifts in turn affect the frequencies of the emitted lines, giving rise to *fine structure*. Attempts to account for the fine structure with the Schroedinger equation lead to failure.

The reader may have studied Einstein's theory of special relativity, which states that the laws of physics must be the same in all reference frames moving at constant velocity with respect to each other, i.e., relativistic invariance. It was known from its inception that the Schroedinger equation does not possess this invariance. It turns out that the simplest way of including special relativity into wave mechanics requires the wave function  $\Psi(x,t)$  of the Schroedinger equation to be replaced by a column matrix with four entries,  $\Psi_1$ ,  $\Psi_2$ ,  $\Psi_3$ ,  $\Psi_4$ , yielding the Dirac equation, a matrix equation. The solutions of the equation make clear that the electron has a spinning motion, the intrinsic spin. This motion has, once a particular axis has been selected, two

senses of rotation, though a unique scalar value of angular momentum.  $\Psi_1, \Psi_2$  denote the relative likelihood of each of the two “spin” possibilities.  $\Psi_3, \Psi_4$  do not correspond to an electron at all, but to a particle of the same mass but opposite charge, the positron, again with two spin possibilities. The discovery of the positron in 1932 confirmed Dirac’s theory, and led to his being awarded the Nobel prize for physics in 1933.

Not only can quantum mechanics, with this modification, account for the fine structure, it can likewise fully account for hyperfine structure. The latter is a tiny shift ( $1/2 \text{ \AA}$ , or less, in the visible region) due to the interaction between the magnetic moment of the nucleus and the magnetic field produced both by the orbital and intrinsic motion of the electron. All these successes, and others in atomic, condensed matter, and nuclear physics establish quantum mechanics as a viable theory over a very broad range of phenomena.

We now turn to a very significant development. In the previous section we saw that the energy levels of the harmonic oscillator are given by the formula:

$$E_n = E_0 + n\hbar\omega \quad n = 1, 2, 3, \dots \quad (6.60)$$

Instead of saying that the oscillator is in its  $n$ th level, we can say that there are  $n$  quanta of oscillation. Such quanta, each with an energy  $\hbar\omega$ , can be thought of as a particle. This new approach, when fully implemented to include other particles, notably electrons and photons, requires for its analysis a further treatment, referred to as *second quantization*. Here, the lowering and raising operators,  $a$  and  $a^+$  of (6.61a,b), become, for obvious reasons annihilation and creation operators, respectively, leading, as they do, to one less or one more quantum.

To bring out second quantization in all its power, we would need to go to a distributed system, for example, the stretched string. We would then treat each mode of oscillation as we are going to treat here the harmonic oscillator, with its one mode of oscillation. The result would be a quantum field theory, such as applies to electromagnetic fields. With some care the theory can be extended to particles that satisfy the Pauli exclusion principle (after the Swiss physicist Wolfgang Pauli, 1900–1958), notably electrons. The principle states that at most one particle per quantum state is allowed. In non-relativistic field theory it is the Schrödinger equation which becomes the equation governing a particle field  $\Psi(r, t)$ , just as the Maxwell equations are the wave equations governing the fields  $E(\mathbf{r}, t)$  and  $B(\mathbf{r}, t)$ . It is all these fields which are quantized, i.e. expressed in terms of the creation and annihilation operators associated with the various modes. The space variables  $r$ , the independent variables in the normal modes are referred to as coordinate parameters. A particle is viewed as one more quantum of oscillation in a given normal mode (eigenstate). The electromagnetic quanta are the photons of Chap. 5. The term second quantization has its origin in as much as for material particles the Schrödinger equation is considered first quantization. While carrying out this development fully is beyond our limited scope here,

we will be able to infer how the different ideas fit together by continuing with our discussion of the Green's function.

In quantum field theory the definition of the dimensionless Green's function for the harmonic oscillator is given by:

$$G(t) \equiv i/(2x_0^2) \int u_0^*(x) [x(t), x(0)] u_0(x) dx \quad t > 0 \quad (6.86a)$$

$$\equiv i/(2x_0^2) \langle u_0 | [x(t), x(0)] | u_0 \rangle \quad t > 0 \quad (6.86b)$$

$$\equiv 0 \quad t \leq 0 \quad (6.86c)$$

Here  $x(t)$  and  $x(0)$  are operators. Their time dependence is generated as in the formula (6.73d) for Heisenberg operators, the transformation operator  $\mathbf{U}(t)$  being given by (6.69). There is no external driving force on the oscillator; the Hamiltonian,  $\mathbf{H}_0$  in (6.49) describes natural motion. We have again taken the initial time  $t_0 = 0$ . The square bracket in (6.86a) denotes the commutator. The scaling parameter  $x_0$  is the length given in (6.78).  $u_0$  is the ground state of the harmonic oscillator. It becomes convenient to no longer regard (6.50a,b) as the defining equations for the creation and annihilation operators, but their inverses, (6.52a,b) as defining the operators  $x$  and  $p$  appearing in (6.86a). Further, when taking expectation values we no longer have to take integrals, but can rely on the orthonormality of the various eigenstates of  $\mathbf{H}_0$  generated by the operators  $a, a^\dagger$ . The variable  $x$ , thus, disappears from this part of the calculation. We adopt the procedure of quantum field theory, and rewrite (6.86a) in the form (6.86b), the triangular brackets again denoting expectation values.

The definition of  $G(t)$  in (6.86a,b) corresponds to what is known in scattering theory as the retarded Green's function, in contrast to an advanced Green's function, see the paper by Carruthers and Nieto listed under Advanced Reading at the end of the chapter. In Chap. 4 we have seen in discussing radiation, that it is the retarded Green's function that can be directly identified with a physical situation. We shall have more to say shortly about the origin of the definitions in (6.86a,b).

We carry out the replacement of the operators  $x(t), x(0)$  by creation and annihilation operators. A moment's reflection will show that the equations (6.50a,b) for  $a, a^\dagger$  can be extended to arbitrary times  $t$  using the formula (6.73d) for conversion from Schrödinger to Heisenberg operators. With the help of (6.52b), we find:

$$[x(t), x(0)] = \frac{\hbar}{2m\omega} [(a^\dagger(t) + a(t)), (a^\dagger(0) + a(0))] \quad (6.87)$$

When we substitute (6.87) for the commutator back into (6.86a) for the Green's function, only two terms survive as we are taking expectation values with the ground state of the oscillator. The careful reader may wish to refer to (6.93a,b). Consistent with the previous section we are here interested in the free oscillator, i.e. without external influence on the propagation. We get:

$$G(t) = i/2 \langle u_0 | \mathbf{a}(t)\mathbf{a}^+(0) - \mathbf{a}(0)\mathbf{a}^+(t) | u_0 \rangle \quad t > 0 \quad (6.88a)$$

$$= 0 \quad t \leq 0 \quad (6.88b)$$

How can we understand the formal expression (6.88a)? In Problem (6.17) the reader is asked to show that an initial wave function,

$$\phi(x, 0) = 1/\sqrt{2} [\mathbf{a}^+(0) - i] u_0(x) \quad (6.89)$$

will lead to the initial expectation values:

$$\overline{x(0)} = 0 \quad (6.90a)$$

$$\overline{p(0)} = \left( \frac{m\hbar\omega}{2} \right)^{1/2} \quad (6.90b)$$

We note that these expectation values correspond to the initial conditions for the quantum Green's function in the previous section, a point to which we shall return. (6.90b) corresponds to (6.83c), with  $x_0$  defined by (6.78), as  $(\hbar/2m\omega)^{1/2}$ . Moreover, the expression obtained for  $x(t)/x_0$  in Problem 6.17 agrees with the full expression (6.88a), obtained from the field theory definition of  $G(t)$ , thereby providing a physical interpretation for (6.88a).

For  $t < 0$ , the oscillator is in the ground state  $u_0$ . Although  $\bar{x}$  is zero, the oscillator is not standing still in this condition. Indeed, as reviewed in Problem 6.16, the uncertainty principle  $\Delta x \Delta p \geq \hbar/2$  is satisfied even here. From Problem 6.17 we learn that the effect of the impulse at  $t = 0$  is a brusque change of state from the ground state to the state described by the normalized function  $1/\sqrt{2} [\mathbf{a}^+(0) - i] u_0(x)$ . We see from this wave function that the oscillator has been jolted from its ground state, with zero average momentum, to a composite state with one component of this state holding one quantum of oscillation. The linear combination of the ground and first excited state leads to the oscillator's continuing to carry, on the average, the acquired momentum. However, a distribution in space and momentum corresponding to the composite wave function is alien to what we find in classical mechanics for a particle. In general, only when large quantum numbers are involved will fully classical features emerge.

The impulse, i.e. momentum  $(m\hbar\omega/2)^{1/2}$  acquired by the oscillator from this impact agrees with the initial impulse acquired by the quantum oscillator in the previous section, i.e., with (6.83c). We note, however, that whereas in the previous section, as in the classical development of Chap. 2, the initial impulse was chosen as a matter of convenience, (fixing the constant in front of  $G$ ), for quantum field theory the impulse is mandated by calculations whose nature is outlined below. The graininess associated with the quantization of the impulse is absent from the two earlier treatments. These we recall, were both based on the model of an external drive, which in the quantum treatment required many quanta in the driving field for its validity. Indeed, the latter calculation is in some contexts referred to as semi-classical; again, we see how, in accord with the correspondence principle, classical features emerge as the number of quanta become large.

As a last step we use what we learned in the previous section about evaluating time dependent operators to find  $G(t)$  as given in (6.88b). Using the formula (6.73d) for Heisenberg operators, and applying (6.69), which gives  $U(t)$  when the Hamiltonian is time independent, we obtain:

$$\mathbf{a}^+(t) = e^{(i/\hbar)\mathbf{H}_0 t} \mathbf{a}^+(0) e^{-(i/\hbar)\mathbf{H}_0 t} \quad (6.91)$$

Applying  $\mathbf{a}^+(t)$  to any one of the eigenstates  $u_n(x)$  we find from (6.91):

$$\mathbf{a}^+(t)u_n(x) = e^{(i/\hbar)\mathbf{H}_0 t} e^{-(i/\hbar)E_n t} \mathbf{a}^+(0)u_n(x) \quad (6.92a)$$

$$= e^{-(i/\hbar)E_n t} e^{(i/\hbar)\mathbf{H}_0 t} (n+1)^{1/2} u_{n+1}(x) \quad (6.92b)$$

$$= e^{-(i/\hbar)E_n t} e^{(i/\hbar)E_{n+1} t} (n+1)^{1/2} u_{n+1}(x) \quad (6.92c)$$

where in (6.92b) we have used (6.61a), which defines the creation operator  $\mathbf{a}^+(0)$ . From (6.92c) we get:

$$\mathbf{a}^+(t)u_n(x) = e^{i\hbar\omega t} \mathbf{a}^+(0)u_n(x) \quad (6.92d)$$

where in the last step we have used (6.60), the formula for  $E_n$ . Since (6.92d) does not depend on  $n$  a similar relation will apply when we operate on any wave function. As a result we have, quite generally:

$$\boxed{\mathbf{a}^+(t) = e^{i\hbar\omega t} \mathbf{a}^+(0)} \quad (6.93a)$$

$$\boxed{\mathbf{a}(t) = e^{-i\hbar\omega t} \mathbf{a}(0)} \quad (6.93b)$$

Here, (6.93b) can be obtained through a similar procedure as that used to obtain (6.93a).

Formulae (6.93a,b), giving the time dependence of the Heisenberg creation and annihilation operators allow us to evaluate (6.88a) for the Green's function. We obtain:

$$G(t) = i/2 (e^{-i\omega t} - e^{+i\omega t}) \langle u_0^* | \mathbf{a}(0) \mathbf{a}^+(0) | u_0 \rangle \quad t > 0 \quad (6.94)$$

and thus:

$$G(t) = 0 \quad t < 0 \quad (6.84a)$$

$$= \sin \omega t \quad t > 0 \quad (6.84b)$$

We have recaptured the results of the previous section. The result for the quantum Green's function, obtained there on the basis of a delta function drive as in classical physics, and the retarded Green's function as defined formally in quantum field theory have been found to agree. Finally, we recall that the wave function of (6.89) which furnished the correct result for the formal expression (6.88a) for  $G(t)$ , also led to the same initial conditions as satisfied by  $G(t)$ , thereby confirming that interpretation of the initial impact.

The question now arises as to the origin of the field theoretic definition of  $G(t)$  in (6.86a,b,c). Two types of problems that use quantum field theory are

particularly prominent in modern physics. The first belongs to the many body problem of solid state physics. We have already indicated that for distributed systems the time dependent part for each normal mode is to be treated as a quantized harmonic oscillator. For example, the collective vibrations of the ions comprising a solid are represented as quantized normal modes, the relevant quanta of energy of vibration being called *phonons*. The interaction of the conduction electrons of the vibrating solids with the phonons is a central subject in many-body theory. The interaction is the mechanism underpinning electrical resistivity in solids on the one hand, and the phenomenon of superconductivity on the other, and of possibly the high temperature superconductivity in the newly developed ceramics as well. When calculations are undertaken depicting this interaction between electrons and phonons, the time evolution of the interacting system gets analysed into a series of apparent (virtual) electron-phonon collisions. In the interval between these collisions the particles propagate, evolve in time, as free. Thus, we need not be surprised that for the phonons these propagators, at their simplest turn out to be Green's functions. The functions depend on an oscillator factor and a space factor. The oscillator contribution conforms with the single oscillator Green's function we have been studying. The Green's functions form some of the building blocks of the overall calculation.

The second type of problem is the basis of *quantum electrodynamics*, the interaction between electrons and the quantized electromagnetic field, i.e. photons. Again, the photon propagators become oscillator Green's functions. What can we say about the electron propagators?

In second quantization electrons can be considered quanta of energy that can be created or annihilated just like photons or phonons. For example, a  $\gamma$  particle, i.e. a high energy photon can, under certain circumstances, split into an electron and its antiparticle, the positron (see below). An electron and a positron have been created (they weren't there before!). The mathematical analysis of the interaction between electron or positron and the photons turns out to picture the evolution in time of the interacting system as a series of virtual collisions. Again one says virtual because such collisions are not what one is setting out to observe; they just appear (not unnaturally) within the mathematical description. In such an analysis one finds that once an electron has been created in a "collision", again the sudden addition of a quantum of energy into the relevant mode, it will, just as a photon, propagate freely until it undergoes the next collision. Evidently, we are dealing with the Green's function for the electron. This function,  $G(2, 1)$  is defined as the probability amplitude for arrival at some point, in one dimension,  $x_2$ , at a time  $t_2$ , starting from  $x_1, t_1$ . The analysis of the electron- positron problem, i.e. pair creation, in terms of Green's functions or propagators was pioneered by the American physicist R.P. Feynman (1918–1988). We are not in a position to write the Feynman propagators in terms of the appropriate creation and annihilation operators and coordinate parameters  $x_1$  and  $x_2$  of second quantization, a task

which would involve a full exposition of quantum electrodynamics. However, assuming that the particle can be treated as non-relativistic, so that the field equation becomes the Schroedinger equation, we do already know the form which  $G(2,1)$  takes. We have from Sect. 6.3, (6.42a,b) with  $k_0 = 0$ :

$$G(2,1) = [a^2/(2\pi^3)]^{1/4} \int_{-\infty}^{\infty} e^{ik(x_2-x_1)} e^{-i\omega_k(t_2-t_1)} e^{-k^2 a^2} dk \quad t_2 \geq t_1 \quad (6.95a)$$

$$= 0 \quad t_2 \leq t_1 \quad (6.95b)$$

$$\omega_k = \hbar^2 k^2 / (2m) \quad (6.42b)$$

Here we are giving the Fourier representation. By integrating over  $k$  we can obtain, for instance, the closed form expression for  $G$  at the initial time  $t_2 = t_1$ , thereby recapturing (6.34b):

$$G(x_2, t_2 = t_1; x_1, t_1) = (2\pi a^2)^{-1/4} e^{-(x_2-x_1)^2/4a} \quad (6.96)$$

The actual propagator  $G(2,1)$  is understood to start off at a point, i.e. one wants to go to the limit  $a \rightarrow 0$ . This is straightforward when we want to use  $G$  as a propagator (Problem 6.18).

We have seen that the origin of the Green's functions as we have defined them in this section is building blocks (propagators) in calculations of quantum field theory. Their definitions in the latter theory can also be in terms of operators whose time dependence is generated by Hamiltonians that include interactions. These more complicated Green's functions contain a wealth of information in their own right. The basic functions, the only ones we have been studying describe unforced or natural propagation. They satisfy Green's function equations as the term was originally understood, and form the basis, historically, of the more elaborate functions. To learn the details of how  $G(1,2)$  in (6.95) satisfies a Green's function equation see Problem 6.18.

Physically the free particle Green's functions, in quantum as well as classical physics depict the situation where the described system is started with a sudden jolt. In quantum mechanics we have seen that such a jolt can be described, for the oscillator, by the sudden addition of a wave function component containing a quantum of oscillation, where previously there were no quanta. Similarly for fermions, particles satisfying the exclusion principle, for example an electron or positron, the jolt occurs when one particle (one field quantum) is created. Second quantization, as, indeed, the whole of quantum theory have their formal origin in a more advanced form of classical mechanics (Lagrangian and Hamiltonian mechanics) that may be unfamiliar to the reader. But even in the more advanced development of quantum theory the Green's function is of help in conceptually bridging the transition from classical mechanics, to pedestrian quantum mechanics, to quantum field theory.

It is quantum field theory that brings about an understanding of the interaction of matter and light. This understanding finds application not

only in such high energy phenomena as electron-positron annihilation, but also, for instance, in the everyday macroscopic application of quantum theory already mentioned, the laser. In Problem 6.20 the interested reader will find discussed the “Lamb shift”, a critical test in its day of validity of quantum electrodynamics (after the American physicist Willis Eugen Lamb, Jr., 1913–).

Einstein’s general relativity postulates an equivalence for the laws of physics no matter what the relative motion of the reference frames. It explains the force of gravity in terms of the curvature of a four dimensional space-time. The forces of nature can be characterized as the electromagnetic forces, the weak forces (that lead to radioactivity), the strong forces that hold together nucleons in the nucleus, and the gravitational forces. While unification of the first three appears to be well on the way, the problem of incorporating general relativity into quantum field theory is not solved. Present efforts are directed towards representing the entities of nature by “strings” in a space-time of possibly twenty-six dimensions. If this theory holds, the development of wave physics will have come full circle, for the dynamics of these elemental strings are taken from those of the classical string under tension in Chap. 3.

**Example 6.1.** Show that the expectation value of the momentum operator  $\bar{p}$  vanishes for the stationary state solutions of a particle entirely confined to a box, as in Fig. 6.1.

*Solution:* From (6.30), a stationary state will have the form

$$\Psi(x, t) = u_m(x)e^{-iE_m t/\hbar}. \quad (6.30)$$

$\bar{p}$  can be found directly from Postulate (3) by

$$\bar{p}(t) = \int_{-\infty}^{\infty} \Psi^*(x, t) \mathbf{p} \Psi(x, t) dx. \quad (6.23a)$$

Substituting for  $\mathbf{p}$  from the table of operators and using (6.30) one gets

$$\begin{aligned} \bar{p}(t) &= \int_{-\infty}^{\infty} e^{+iE_m t/\hbar} e^{-iE_m t/\hbar} u_m^*(x) (\hbar/i) \partial u_m(x) / \partial x dx \\ &= \int_{-\infty}^{\infty} u_m^*(x) (\hbar/i) \partial u_m(x) / \partial x dx. \end{aligned}$$

From (6.14)

$$\begin{aligned} u_m(x) &= \sqrt{2/L} \sin(m\pi x/L) \\ &= \frac{1}{\sqrt{2L}} \frac{1}{i} (e^{ik_m x} - e^{-ik_m x}) \end{aligned} \quad (6.14)$$

with  $k_m = m\pi/L$ .

Thus,

$$\bar{p}(t) = \frac{L}{2} \int_0^L (e^{+ik_m x} - e^{-ik_m x})^* \left( \frac{\hbar}{i} \frac{\partial}{\partial x} \right) (e^{ik_m x} - e^{-ik_m x}) dx .$$

We note that

$$\begin{aligned} \int_0^L (e^{ik_m x})^* e^{-ik_m x} dx &= [i/2k_m] (e^{-2ik_m L} - 1) \\ &= i/(2k_m) (e^{-i2\pi m} - 1) = 0 , \end{aligned}$$

(the two functions are orthogonal).

Similarly,

$$\int_0^L (e^{-ik_m x})^* e^{ik_m x} dx = 0 .$$

We are left with

$$\begin{aligned} \bar{p}(t) &= -(L/2) \int_0^L [e^{-ik_m x} \hbar k_m e^{ik_m x} + e^{ik_m x} (-\hbar k_m) e^{-ik_m x}] dx \\ &= 0 . \end{aligned}$$

**Example 6.2.** Find the eigenstates and energies of a particle confined by a delta function potential well given by

$$V(x) = -V_0 \delta(x) ,$$

where  $V_0$  is a positive constant (with dimensions of energy times length). Delta function potentials are often used as models in quantum mechanics. This example will be used in Chap. 7, on *Nonlinear Waves – Solitons*.

*Solution.* We are to solve the Schrödinger time-independent equation

$$-(\hbar^2/2m)d^2/dx^2 + V(x)u(x) = Eu(x) \quad (6.3c)$$

where  $V(x) = -V_0 \delta(x)$ .

Simplifying, we may write (6.3c) as

$$\frac{d^2 u}{dx^2} + [\lambda + W\delta(x)]u(x) = 0$$

where  $W = (2m/\hbar^2)V_0 > 0$  and  $\lambda = (2m/\hbar^2)E$ . Moreover, since the particle is bound, i.e., confined to the neighborhood of  $x = 0$ , we expect that whatever its value  $E$ , the total energy and, hence,  $\lambda$ , are negative;  $\lambda < 0$ . Outside of  $x = 0$ , that is, outside the well,  $u(x)$  must satisfy the simpler equation

$$\frac{d^2u}{dx^2} + \lambda u(x) = 0, \quad \lambda < 0.$$

The only possible solutions here are

$$\begin{aligned} u(x) &= c_1 e^{-Kx} & x > 0 \\ u(x) &= c_2 e^{Kx} & x < 0 \\ K^2 &= -\lambda, & K > 0 \end{aligned}$$

We have exponentials of real functions. We choose decaying functions as  $x \rightarrow \pm\infty$  to allow normalization of the wave function. Further, continuity of  $u(x)$  requires  $c_1 = c_2 \equiv c$ . Integration of the Schrödinger equation across the delta function potential (see also Problem 6.11) yields

$$\left. \frac{du}{dx} \right|_{0+} - \left. \frac{du}{dx} \right|_{0-} = -W u(0) = -Wc.$$

Here, the well extending from  $0^+$  to  $0^-$ , is recognized to have vanishing width, so that the integral of the term  $\lambda u(x)$  across the well vanishes. From the above equation, substituting the given solutions for  $u(x)$  we find

$$-Kc - Kc = -Wc, \quad K = W/2.$$

Thus we find

$$\begin{aligned} u(x) &= c e^{-Wx/2} & x > 0 \\ u(x) &= c e^{Wx/2} & x < 0 \end{aligned}$$

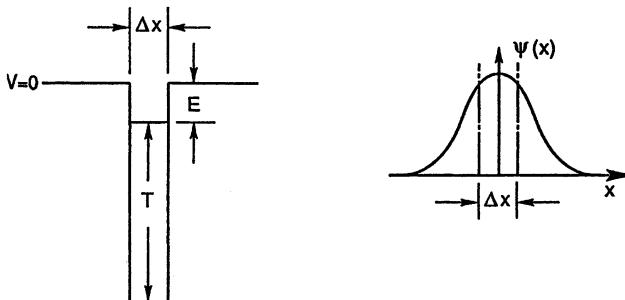
and  $E = \hbar^2 W^2 / 8m$ . Normalization yields

$$\int_{-\infty}^{\infty} u^*(x)u(x)dx = 2 |c|^2 \int_0^{\infty} e^{-Wx} dx = 1$$

or  $|c| = \sqrt{W/2}$ .

We see that the solution is unique. The delta function potential in one dimension has always exactly one bound state. In three dimensions, it turns out that a certain width is needed to secure even one bound state, see the text by L. I. Schiff listed under Further Reading.

According to classical physics the particle would be confined to within the well; its kinetic energy is less than  $V_0$ , since its total energy  $E$  has been chosen to be negative. This means the particle does not have enough energy to climb over the potential barrier at each side of the potential well. In quantum mechanics we do find a certain probability of its being outside. Actually, with a delta function potential the particle is always outside, but because of its total negative energy it cannot leave the well and become free, i.e., have an imaginary exponential solution  $e^{ikx}$ , with  $k$  real. If the particle were confined inside,  $\Delta x$  would go to zero, which would cause  $\Delta k \rightarrow \infty$ , implying infinite kinetic energy (Fig. 6.3). Exponentially decaying solutions occur frequently in quantum mechanics, particles “tunnel” through so-called potential barriers, which are regions of negative kinetic energy for the particles, regions which



**Fig. 6.3.** Example 6.2: Particle confined to a nearly  $\delta$  function potential, shown on the *left*. Inside, the kinetic energy  $T > 0$ . Outside, its kinetic energy becomes its total energy,  $T < 0$ . On the *right* the wave function for the particle

in classical mechanics totally repel. Radioactive decay of nuclei, common in nature, takes place by means of such tunneling processes. Similarly the remarkable superconductive *Josephson junctions* involve electron tunneling.

**Example 6.3.** Show that visual inspection can not locate a particle and determine the momentum simultaneously to better than

$$\Delta x \Delta p \sim h .$$

*Hint:* Recall the effect of diffraction on resolution.

*Solution.* We recall from Chap. 5 that the angle of resolution  $\theta_R$  is given approximately by

$$\theta_R \simeq \lambda/D ,$$

where  $D$  is the lens diameter (Fig. 6.4). Diffraction, therefore, restricts our ability to locate the particle to within an uncertainty

$$\Delta x \gtrsim \theta_R d ,$$

$$\gtrsim \lambda d/D ,$$

where  $d$  is the distance between lens and particle. The particle is located by photons bouncing from it. The momentum carried by a photon  $p$  is given by

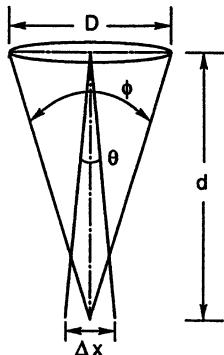
$$p = h/\lambda .$$

We do not know which path within the angle  $\phi$  the photon followed. The particle, we note, experiences a recoil, leaving us with an uncertainty  $\Delta p$  as to its final momentum of

$$\Delta p \gtrsim (h/\lambda)\phi \gtrsim (h/\lambda)D/d .$$

Thus, combining the above results for  $\Delta x, \Delta p$  we get

$$\Delta x \Delta p \gtrsim h .$$



**Fig. 6.4.** Example 6.3. Finite resolution of a viewing lens leads to an uncertainty in position  $\Delta x$ . The reflected photon may have traveled anywhere within the cone subtended by the angle  $\phi$

**Example 6.4.** Consider a system, described by a Hamiltonian  $H_0$ , which for practical purposes has only two eigenstates,  $u_1, u_2$ , with energies  $\hbar\omega_1, \hbar\omega_2$ ,  $\omega_2 > \omega_1$ . These eigenstates could correspond, for instance, to the situation of a nuclear magnetic moment  $\mu$  being parallel or antiparallel to a constant (D.C.) magnetic field  $B_0$ , leading to an energy splitting  $2\mu B_0$  between the two states. In this case we would have  $2\mu B_0 = (\hbar\omega_2 - \hbar\omega_1)$ . Suppose now, that this system is driven from the outside harmonically with a driving frequency  $\omega$ , adding a term  $\mathbf{W} \sin \omega t$  to the Hamiltonian, which becomes:

$$\mathbf{H} = \mathbf{H}_0 + \mathbf{W} \sin \omega t$$

In the case of the nucleus in the field  $B_0$ , the new term could represent the influence of an additional small time dependent field  $\mathbf{B}(t) = \mathbf{B} \sin \omega t$ , applied at right angles to  $B_0$ . In classical physics, such a field would exert a torque  $\mu \times \mathbf{B}(t)$ , tending to rotate the nucleus relative to  $B_0$ . Quantum mechanically, the time dependent field would introduce the term  $\mathbf{W} \sin \omega t$  into  $\mathbf{H}$ ,  $\mathbf{W}$  being proportional to the product of  $\mu$  and  $B(t)$ . This term would lead to the transitions between the two “spin” (up and down) states. (The expression for  $W$  requires a knowledge of the Pauli spin matrices.)

In our problem, generally, we may assume that the overlap integrals

$$W_{12} = \int u_1^*(x) \mathbf{W} \sin \omega t u_2(x) dx$$

and its adjunct  $W_{21}$  are real and equal. However, they may not vanish. For the spinning nucleus, the integral would be replaced by a sum over the possible values of the spin (up or down relative to  $B_0$ , with  $\mathbf{W}$  connecting the spin states).

a) On the assumption that the system is entirely in the lower energy states  $u_1$ , at  $t = 0$ , find the probability that the system will be found in state  $u_2$  at some subsequent time  $t$ . Discuss the appearance of resonance. Assume throughout that the driving frequency  $\omega$  is close to the energy splitting of states (1) and (2):

$$|\omega - (\omega_2 - \omega_1)| \ll \omega_2 - \omega_1, \quad \omega_2 > \omega_1.$$

b) Find a formula for the time average response when the system is a spinning nucleus under the influence of the D.C. magnetic field  $\mathbf{B}_0$ , and the driving A.C. magnetic field  $\mathbf{B} \sin \omega t$ .

*Solution:* As suggested in Sect. 6.5, we expand the time dependent solution  $\Psi(t)$  as:

$$\Psi(t) = \sum_{i=1}^2 c_i(t) u_i e^{-i\omega_i t}$$

where

$$\mathbf{H}_0 u_i = \hbar \omega_i u_i \quad i = 1, 2.$$

Substituting into the time-dependent equation

$$i\hbar \partial \Psi(t) / \partial t = (\mathbf{H}_0 + \mathbf{W} \sin \omega t) \Psi(t)$$

we find:

$$\begin{aligned} & i\hbar \sum_i [dc_i/dt u_i - i\omega_i c_i u_i] e^{-i\omega_i t} \\ &= \sum_i [c_i(t) \hbar \omega_i u_i + \mathbf{W} \sin \omega t c_i(t) u_i] e^{-i\omega_i t} \end{aligned}$$

or, cancelling the biggest terms,

$$i\hbar \sum_i dc_i/dt u_i e^{-i\omega_i t} = \sum_i c_i(t) \mathbf{W} \sin(\omega t) u_i e^{-i\omega_i t}$$

If we multiply by  $u_1(x)$  and integrate over  $x$ , (or sum over spins), we obtain, recalling that the  $u$ 's are orthonormal:

$$i\hbar (dc_1/dt) e^{-i\omega_1 t} = c_2(t) W_{12} (\sin \omega t) e^{-i\omega_2 t}$$

On the assumption that  $W_{11} = W_{22} = 0$ ,  $W_{12} = W_{21} \equiv \hbar w$  we obtain:

$$-2\hbar dc_1/dt = \hbar w c_2 \left( e^{i[\omega - (\omega_2 - \omega_1)]t} - e^{-i[\omega + (\omega_2 - \omega_1)]t} \right)$$

We may ignore the second exponential, on the assumption that this term has such a rapid time variation as compared to the first exponential, that on the average over time it will not contribute. On the other hand, we recall that

$$\Delta\omega \equiv \omega - (\omega_2 - \omega_1)$$

the difference between the driving frequency, and the frequency split of our two eigenstates, is to be considered small. As a result we get:

$$-2 \frac{dc_1}{dt} = w c_2 e^{i\Delta\omega t}$$

Proceeding similarly, but first multiplying by  $u_2(x)$  we obtain

$$-2 \frac{dc_2}{dt} = -c_1 e^{-i\Delta\omega t}$$

After noting the asymmetry of the last two equations, we look for a normal mode solution to the coupled linear differential equations by letting:

$$\begin{aligned}c_1(t) &= a_1 e^{i\Omega t} e^{i\Delta\omega t/2} \\c_2(t) &= a_2 e^{i\Omega t} e^{-i\Delta\omega t/2}\end{aligned}$$

This substitution yields:

$$\begin{aligned}-2a_1 i(\Omega + \Delta\omega/2) &= wa_2 \\-2a_2 i(\Omega - \Delta\omega/2) &= -wa_1\end{aligned}$$

As usual, (see Problems 3.9 and 6.14) we obtain the “secular equation”:

$$\begin{vmatrix} -i(2\Omega + \Delta\omega) - w \\ w - i(2\Omega - \Delta\omega) \end{vmatrix} = 0$$

or,

$$\pm\Omega = \sqrt{(\Delta\omega)^2 + w^2}/2; \quad \Omega > 0.$$

Thus, we obtain

$$c_1(t) = e^{i\Delta\omega t/2} (A e^{i\Omega t} + B e^{-i\Omega t})$$

where A and B are the amplitudes of the two possible “normal modes”, each mode having its respective frequency. Further, we have available from our solution of the coupled linear equations the ratios  $a_2/a_1$  for each of the two frequencies,  $\pm\Omega$ . Thus:

$$c_2(t) = -ie^{-i\Delta\omega t/2} \frac{w}{2} \left( \frac{A e^{i\Omega t}}{\Omega - \Delta\omega/2} - \frac{B e^{-i\Omega t}}{\Omega + \Delta\omega/2} \right)$$

At  $t = 0$ ,  $c_1(t) = 1$ ,  $c_2(t) = 0$ , the given initial conditions, normalization provided. It follows that:

$$A = (\Omega - \Delta\omega/2)/(2\Omega) \quad B = (\Omega + \Delta\omega/2)/(2\Omega)$$

Substituting these values for A and B back into our expressions for  $c_1(t)$  and  $c_2(t)$  we obtain:

$$\begin{aligned}c_1(t) &= e^{i\Delta\omega t/2} [\cos \Omega t - i\Delta\omega/(2\Omega) \sin \Omega t] \\c_2(t) &= e^{-i\Delta\omega t/2} w/(2\Omega) \sin \Omega t\end{aligned}$$

a) We are asked to find the probability that the system is to be found in state  $u_2$  at a subsequent time  $t$ , assuming, as we have, that the system is entirely in state  $u_1$  at  $t = 0$ , i.e.  $c_1(0) = 1, c_2(0) = 0$ . Clearly, this probability,  $P(t)$  is given by:

$$\begin{aligned}P(t) &= |c_2(t)|^2 = \frac{w^2}{(\Delta\omega)^2 + w^2} \sin^2 \Omega t \\4\Omega^2 &= (\Delta\omega)^2 + w^2 \quad W_{12} = W_{21} \equiv \hbar w\end{aligned}$$

This result shows that the system alternates as  $\sin^2 \Omega t$  between the two eigenstates,  $u_1$  and  $u_2$ .

- b) The “response” of a spinning nucleus, i.e. the time average probability of finding the upper state occupied, with the corresponding average energy drawn from the transverse driving field  $B_0 \sin \omega t$  is, from Part (a)

$$\overline{|c_2(t)|^2} = w^2 / (\Delta^2 \omega + w^2)$$

This is again a resonance response, reaching its maximum when  $\Delta\omega$  vanishes, i.e. as

$$\Delta\omega \equiv \omega - (\omega_2 - \omega_1) \rightarrow 0$$

Here  $\omega$  is the driving frequency, and  $(\omega_2 - \omega_1)$ , the splitting of the two states, the “natural frequency” of the system. We have seen that  $w$ , i.e. overlap integrals  $W_{12}, W_{21}$  is proportional to the driving magnetic field  $B$ . By choosing  $B \ll B_0$ , ( $\hbar(\omega_2 - \omega_1) = 2\mu B_0$ ), one can obtain a sharp resonance.

We are back to driven motion as distinct from natural motion. Indeed, we find all the connections between the two types of motion as we did in the earlier chapters on classical physics. In Problem 6.19 the external agent is an alternating electric field, and we find an electric susceptibility that shows the same characteristics at resonance as yielded by the representation of an atom as a classical spring in Problem 2.10. The natural frequency of the spring is replaced by  $\Delta E/\hbar$ , where  $\Delta E$  is again the energy spacing of the two stationary states. We do need to note, however, that there are also differences with the classical, for it is the expectation values of physical quantities, rather than the wave function  $\Psi$ , that respond at the driving frequency. A perusal of Problem 6.13 will be useful in gaining additional insight into how it is  $\Psi^* \Psi$  rather than  $\Psi$  that can relate to phenomena observed in classical physics.

Example 6.4 with the energy splitting of the two states coming as a consequence of the interaction between the angular momentum of a nucleon and an applied D.C. magnetic field underlies magnetic imaging techniques in medicine. If the second (the A.C.) magnetic field is applied as a pulse, the resulting resonance condition quickly leads to equal occupation of both stationary states. The times for the excitation to return to equilibrium once the pulse is over determine the induced output. These relaxation times depend on the magnetic fields in the microscopic neighborhood of each resonating nucleus, with different fields in different tissues.

In conclusion, we have found that there are strong analogues between classical and quantum mechanical responses. In the first instance oscillators are associated with energy eigenstates just as there are oscillators associated with normal modes; it is the treatments of the oscillators which run on parallel if not identical lines.

## Problems

**6.1** a) An electron and a photon each have a wavelength of 2.0 Å. What are

- i) their momenta?
- ii) their energies?

b) Assuming equal lens diameters, how does the resolving power of a 10 keV electron microscope compare to an ordinary microscope using light?

**6.2** Verify that the method of separation of variables, i.e., a trial solution of the form

$$\Psi(x, t) = u(x)v(t)$$

indeed leads to a solution of Schroedinger's equation (6.4), with  $u(x)$  satisfying (6.3), and  $v(t)$  ultimately satisfying (6.2a) and (6.6).

**6.3** a) Calculate the Bohr radius  $a$  and the total energy  $E_1$  of an electron in the ground state of a hydrogen atom, using Bohr's semiclassical theory.

*Hint:* To find the radius  $a$ , equate the centripetal acceleration  $mv^2/a$  to the centripetal force  $e^2/(4\pi\varepsilon_0a^2)$ , while fixing the electron's angular momentum  $mva$  at  $n\hbar$ , with  $n = 1$  ( $e$  is the charge of the electron). Then show that the total energy  $E_1$  is given by  $-e^2/(8\pi\varepsilon_0a)$ , and substitute in the expression for  $a$ .

b) The time-independent wave function for the ground (1s) state is given by

$$u_1(r) = (\pi a^3)^{-1/2} e^{-r/a} .$$

Use the time-independent Schroedinger equation

$$[-(\hbar^2/2m)\nabla^2 + V(\mathbf{r})]u(\mathbf{r}) = Eu(\mathbf{r})$$

to show that  $a$  and  $E_1$  are uniquely given by the same expressions as in the semiclassical Bohr theory above.

*Hints:* For the spherically symmetric solution  $u_1(r)$  the Schroedinger equation becomes

$$\left[ -\frac{\hbar^2}{2mr^2} \frac{\partial}{\partial r} \left( \frac{r^2\partial}{\partial r} - \frac{e^2}{4\pi\varepsilon_0 r} \right) \right] u_1(r) = E_1 u_1(r) .$$

The two sides of the equation must agree for all values of  $r$ .

**6.4** a) Verify the derivation of (6.24b) from the definition of  $\overline{(\mathbf{Q} - \bar{\mathbf{Q}})^2}$ .

b) Calculate  $\Delta p$  for a particle totally confined to a potential well,  $L > x > 0$  (Fig. 6.1) in the state  $u_n(x)$ , with  $n = 3$ .

c) Calculate  $\bar{x}$  and  $\Delta x$  for the same eigenstate as in (b). **6.5 Periodic Boundary Conditions.** Consider the potential well with

$$\begin{aligned} V = 0 &\quad L > x > 0 \\ V = \infty &\quad x > L, \quad x < 0 . \end{aligned}$$

If  $L$  is large compared to the region of interest in a particular problem, we can impose periodic boundary conditions:

$$\Psi(x + L) = \Psi(x).$$

That is, as regards boundary conditions, the straight line interval  $x = 0$  to  $x = L$  is replaced by a circle of circumference  $L$ . Since the boundaries are assumed to have no influence this physical model is both convenient and acceptable.

- a) Write down all possible time-independent solutions within the well. Normalize these solutions.
- b) Show that your solutions are eigenstates of both the total Hamiltonian and of the momentum operator.
- c) Show that for any energy eigenvalue less than or equal to some eigenvalues  $E_n$  of the periodic solution there are as many standing wave solutions  $u(L) = u(0) = 0$  as there are periodic solutions.

**6.6** a) Given the wave number distribution (6.35b) for a free particle show that  $\Delta p(0) = \hbar/2a$ .

*Hint:* Refer to the general formulas (6.24b, 28b) and the derivation of (6.33a) for  $\bar{p}$ .

- b) Similarly, noting (6.39), show that for the same initial state  $\Delta x(0) = a$ .

**6.7 Commuting Operators.** The commutator of two operators  $A, B$ , namely  $[A, B]$ , is defined as

$$[A, B] = AB - BA.$$

- a) Show that operators that commute have the very same eigenfunctions.
- b) Calculate

i)	$[\mathbf{p}, x]$	$\mathbf{p} = (\hbar/i)\partial/\partial x$
ii)	$[\mathbf{p}, \mathbf{H}]$	$\mathbf{H} = V(x) + T$
iii)	$[\mathbf{p}, \mathbf{T}]$	$V(x) \neq 0$
		$\mathbf{T} = -(\hbar^2/2m)\partial^2/\partial x^2$ .

- c) On the basis of parts (a) and (b) suggest why it might be possible to measure the momentum and energy of a *free* particle simultaneously without error, whereas the Heisenberg uncertainty principle limits a simultaneous determination of, say, the momentum and position.

**6.8** A certain telescope allows the observer to locate the moon within  $5 \times 10^{-7}$  radians. How many seconds will it take for the moon to deviate by 1 km from its most probable course with a reasonable likelihood? Assume a free particle model.

**6.9** a) Consider a one dimensional square well extending from  $x = -a$  to  $x = a$ , with constant negative potential  $-V_0$ .

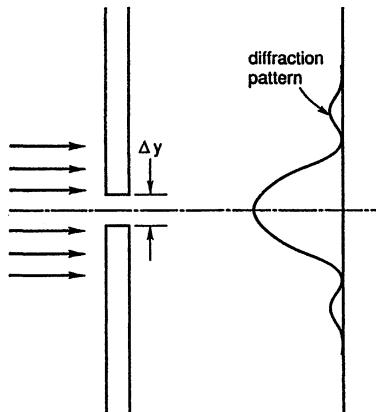
$$\begin{aligned} V(x) &= -V_0, & V_0 > 0, & a > x > -a \\ V(x) &= 0 & x < -a, & x > a. \end{aligned}$$

Show that no matter how small the quantity  $V_0a^2$  is, there is always at least one bound state with  $E < 0$ . What is the time-independent wave function for this state  $u_1(x)$ ?

*Hint:* See Fig. 6.3.

b) The reflection operator  $\mathbf{R}$  changes any function  $u(x)$  or operator  $Q(x)$  into  $u(-x)$ ,  $Q(-x)$ , respectively. Show that for the Hamiltonian of Problem 6.9(a),  $\mathbf{H}$  and  $\mathbf{R}$  commute (Problem 6.7). This implies that the solutions of (a) have definite symmetry about  $x = 0$ ; explain. What symmetry does the solution  $u_1(x)$  that you found in (a) have?

**6.10** Show that if a single horizontal slit in a baffle is used to fix the location of a particle as it crosses the baffle (Fig. 6.5), the simultaneous vertical momentum,  $p_y$  and location,  $\Delta y$ , of the particle can again only be inferred to an accuracy within the Heisenberg uncertainty.



**Fig. 6.5.** Electron beam diffracted by a single slit, Problem 6.10

**6.11** Arrange a table as shown, numbering slots; then fill in the called for developments.

Show that  $T$  the transmission probability across a sudden discontinuity in the density of a long string (Fig. 6.6a) is given by

$$T = \left(2\sqrt{\varrho_1}/(\sqrt{\varrho_1} + \sqrt{\varrho_2})\right)^2 \sqrt{\frac{\varrho_1}{\varrho_2}}.$$

Here  $\varrho_1, \varrho_2$  are the densities to the right and left, respectively, of the discontinuity.  $T$  is defined by

$$T \equiv (C/A)^2 v_2/v_1$$

(See below.)

Show that  $T$  the transmission probability across a sudden potential barrier, (Fig. 6.6b) is given by

$$T = \left(2\sqrt{E}/(\sqrt{E} + \sqrt{E - V_0})\right)^2 \sqrt{\frac{E - V_0}{E}}.$$

Here  $E$  is the total energy of the particles,  $V_0$  the barrier “height”.  $T$  is defined by

$$T \equiv (C/A)^2 (v_2/v_1) \\ (v \equiv \hbar k/m)$$

(See below.)

- a) Write the time-dependent wave equation for  $y_1(x, t)$ , the displacement for  $x < 0$ .

- b) Make the substitution

$$y_1(x, t) = \operatorname{Re}\{u_1(x)e^{-i\omega t}\} \text{ for } x < 0.$$

Substitute into (a), and find the time-independent wave equation for  $u_1(x)$ .

- c) Repeat (a) and (b) with

$$y_2(x, t) = \operatorname{Re}\{u_2(x)e^{-i\omega' t}\} \text{ for } x > 0.$$

- d) Let  $u_1(x) = Ae^{ik_1 x} + Be^{-ik_1 x}$   
and  $u_2(x) = Ce^{ik_2 x}$

Discuss.

Find  $k_1$  and  $k_2$  in terms of  $\omega$  and  $\omega'$ .

- e) Show on physical grounds that

$$\lim_{x \rightarrow 0^-} y_1(x, t) = \lim_{x \rightarrow 0^+} y_2(x, t)$$

- f) Use (e) to obtain a relation between  $\omega$  and  $\omega'$ .

- g) Use Newton's 2nd law to show that  $\partial u / \partial x$  is continuous at  $x = 0$  unless  $\varrho = \infty$ .

*Hint:* go back to derivation of the wave equation in Chap. 3.

- h) Using (d – g) show that

$$(C/A)^2 = \left( \frac{2k_1}{k_1 + k_2} \right)^2.$$

- i) Explain the definition of  $T$  given at the beginning, and verify the answers.

- a) Write the time-dependent wave equation for  $\Psi_1(x, t)$ , the wave function for  $x < 0$ .

- b) Make the substitution

$$\begin{aligned} \Psi_1(x, t) &= u_1(x)e^{-i(E/\hbar)t} \\ &\equiv u_1(x)e^{-i\omega t} \text{ for } x < 0. \end{aligned}$$

Substitute into (a) and find the time-independent equation for  $u_1(x)$ .

- c) Repeat (a) and (b) with

$$\begin{aligned} \Psi_2(x, t) &= u_2(x)e^{-i(E'/\hbar)t} \\ &\equiv u_2(x)e^{-i\omega' t} \text{ for } x > 0. \end{aligned}$$

- d) Let  $u_1(x)Be^{-ik_1 x}$   
and  $u_2(x)Ce^{ik_2 x}$

Discuss.

Find  $k_1$  and  $k_2$  in terms of  $E$  and  $E'$ .

- e) The continuity of  $\Psi$  is one of the postulates of quantum mechanics.  
Show that

$$\lim_{x \rightarrow 0^-} \Psi_1(x, t) = \lim_{x \rightarrow 0^+} \Psi_2(x, t)$$

- f) Give  $\omega$  and  $\omega'$  in terms of  $E$  and  $E'$ .

- i) Use (e) to obtain a relation between  $\omega$  and  $\omega'$ .

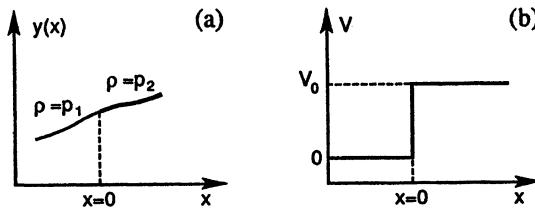
- ii) What is the relation between  $E$  and  $E'$ ? Do you believe it?

- g) Use the general form of the time-dependent equation obtained in (b) to show that  $\partial u / \partial x$  is continuous at  $x = 0$ , indeed, always, unless  $V = \infty$ .

- h) Using (d – g) show that

$$(C/A)^2 = \left( \frac{2k_1}{k_1 + k_2} \right)^2.$$

- i) Explain the definition of  $T$  given at the beginning, and verify the answers.



**Fig. 6.6.** (a) Stretched string with a discontinuity in its mass per unit length at  $x = 0$ . (b) The corresponding discontinuity in the potential  $V(x)$ , Problem 6.11

**6.12** A particle of total energy  $E$  tunnels through a potential barrier  $V(x)$  from  $x_0$  to  $x_f$ , where  $V(x) > E$ ,  $x_f > x > x_0$ . Show that if in one wavelength the fractional change of potential energy as compared to its kinetic energy is small, i.e.,

$$\frac{|\,dV(x)/dx\,| \lambda}{(V - E)} \ll 1,$$

then the probability of the particle's getting through the barrier is given by

$$\exp \left\{ - \left( 2^{3/2} m^{1/2} / \hbar \right) \int_{x_0}^{x_f} [V(x) - E]^{1/2} dx \right\}.$$

**6.13** a) We define a vector  $\mathbf{S}(r, t)$  by

$$\mathbf{S} = \frac{\hbar}{2im} [\Psi^* \nabla \Psi - (\nabla \Psi^*) \Psi].$$

Use Schroedinger's time-dependent equation and integration by parts to show that

$$\frac{\partial}{\partial t} [\Psi^*(r, t) \Psi(r, t)] + \nabla \cdot \mathbf{S}(r, t) = 0.$$

b) Give arguments to support the interpretation of  $N\mathbf{S}$  as a current density, where  $N$  is the number of particles present per unit volume.

c) Working in one dimension only show that

$$m \frac{d\bar{x}}{dt} = \frac{\hbar}{2i} \int_{-\infty}^{\infty} \left( \Psi^* \frac{\partial \Psi}{\partial x} - \frac{\partial \Psi^*}{\partial x} \Psi \right) dx.$$

Argue that the velocity of a particle  $v$  ( $\equiv d\bar{x}/dt$ ) must satisfy  $S_x(x, t) = v \Psi^*(x, t) \Psi(x, t)$ , if  $v$  is constant over the spread of the particle. Use parts (a) and (c) to verify this relation. Further, since  $\bar{\mathbf{p}}_x$  is always real, show that

$$\frac{d\bar{x}}{dt} = \frac{1}{m} \bar{\mathbf{p}}_x.$$

d) Again integrating twice by parts, use Schroedinger's equation to verify that

$$\frac{d\bar{\mathbf{p}}_x}{dt} = -\frac{\partial \bar{V}}{\partial x} = -\int \Psi^* \frac{\partial V}{\partial x} \Psi dx.$$

Recalling (c), show that Newton's second law is regained if  $-\partial V/\partial x$ , the force, varies little over the extent of a wave packet describing a particle (Ehrenfest's Theorem).

**6.14** To illustrate a technique for solving time-dependent problems in wave mechanics we consider a system with only two degenerate eigenstates  $u_0(x)e^{-iEt/\hbar}$  and  $u_1(x)e^{-iEt/\hbar}$ . Degenerate means that both states have the same energy  $E$ , although the states are orthogonal. Thus:

$$\mathbf{H}_0 u_0(x) = E u_0(x), \quad \mathbf{H}_0 u_1(x) = E u_1(x)$$

$u_0$  and  $u_1$  are normalized, and orthogonal to each other. Suppose that at time  $t = 0$ , when the system is in one of the states,  $u_0(x)$ , it is disturbed by an external agent,  $W(x)$ , so that the Hamiltonian becomes:

$$\mathbf{H} = \mathbf{H}_0 + W(x).$$

Find the resulting time-dependent function  $\Psi(x, t)$  after the disturbance. Discuss the new physical situation, comparing your solution with the situation encountered for two lightly coupled harmonic oscillators as in Problem 3.9.

*Hints:* Let  $\Psi(x, t) = [a_0 u_0(x) + a_1 u_1(x)]e^{-i\Omega t}$  where  $\hbar\Omega$  is the energy of new stationary states which include the effect of  $W(x)$ . Obtain the secular determinant

$$\begin{aligned} & \left| \begin{array}{l} -\hbar\Omega + E \quad W_{01} \\ W_{01}^* - \hbar\Omega + E \end{array} \right| = 0 \\ & W_{01} = \int_{-L/2}^{+L/2} u_0^*(x) W(x) u_1(x) dx \quad \text{Does } W_{01} = W_{10}^*? \end{aligned}$$

Now, let  $\Psi(x, t) = \alpha_1 \phi_1(x, t) + \alpha_2 \phi_2(x, t)$  where  $\phi_1, \phi_2$  are the new eigenstates, and  $\alpha_1, \alpha_2$  are to be determined from the initial condition. Finally, find  $A_0(t)$  and  $A_1(t)$ , defined by  $\Psi(x, t) \equiv A_0(t)u_0(x) + A_1(t)u_1(t)$ .

**6.15** a) Use (6.51) to show that the eigenvalues of a Hermitian operator are real.

*Hint:* Expand  $v(x)$  and  $w(x)$  in the eigenfunctions of the operator, and, at some point recall that  $v$  and  $w$  are, if arbitrary, yet physical (good) functions. Can you also argue in the reverse direction, i.e. prove sufficiency?

b) Prove (6.71).

*Hint:* Show first that:

$$\int_{-\infty}^{\infty} \left[ \frac{d}{dx} v(x) \right]^* w(x) dx = - \int_{-\infty}^{\infty} v^*(x) \frac{d}{dx} w(x) dx$$

c) Show that the operator  $(i\mathbf{H})^n$  satisfies (6.70a) viz,  $\{(i\mathbf{H})^n\}^+ = \{(i\mathbf{H})^n\}^*$ .

*Hint:* Show first that  $\mathbf{H}^n$  is Hermitian if  $\mathbf{H}$  and  $(\mathbf{H})^{n-1}$  are,  $n = 3, 4, \dots$ ; how about if  $n = 2$ ?

**6.16** a) Show by direct substitution that the function

$$u_0(x) \equiv (2\pi x_0^2)^{-1/4} \exp^{-[x^2/(4x_0^2)]}$$

is the solution of Schroedinger's time independent equation:

$$\mathbf{H}_0 u_0(x) = (\hbar\omega/2)u_0(x),$$

$\mathbf{H}_0$  given by (6.49).

b) Find, using (6.50a)  $\mathbf{a}u_0(x)$ ,  $\mathbf{a}$  the annihilation operator.

c) Find  $\Delta x, \Delta p$ , and evaluate the product  $\Delta x \Delta p$ , for the function  $u_0(x)$ , as given above.

**6.17** a) Evaluate  $\bar{x}(0)$  and  $\bar{p}(0)$ , expectation values at  $t = 0$  for the initial wave function

$$\phi(x, 0) = 1/\sqrt{2} [\mathbf{a}^+(0) - i] u_0(x).$$

$u_0(x)$  is the ground state for the harmonic oscillator.

*Hint:* Refer to (6.52a,b)

b) Show that, in agreement with (6.88a),  $\bar{x}(t)$  the expectation value at time  $t$ ,  $t > 0$ , is given by:

$$\begin{aligned}\bar{x}(t) &= ix_0/2 \langle u_0 | \{\mathbf{a}(t)\mathbf{a}^+(t) - \mathbf{a}(0)\mathbf{a}^+(t)\} | u_0 \rangle \\ x_0^2 &\equiv \hbar/(2m\omega)\end{aligned}$$

**6.18** a) Carry out the  $k$  integration in (6.95a), and compare your result with (6.42e) and (6.42h) in Sect. 6.3

b) Show that  $G(2, 1)$  as given in (6.95) propagates, i.e.

$$G(3, 1) = \int_{-\infty}^{\infty} G(3, 2)G(2, 1)dx_2$$

*Hint:* Use the Fourier representation (6.95). Seek a  $\delta$  function representation for the two wave-vectors in the representation.

c) Show that  $G(2, 1)$  in (6.95a,b) satisfies the Green's function equation:

$$\left( i\hbar \frac{\partial}{\partial t_2} - \frac{\hbar^2}{2m} \frac{\partial^2}{\partial x_2^2} \right) G(2, 1) = i\hbar(2\pi a^2)^{-1/4} e^{-(x_2-x_1)^2/(4a^2)} \delta(t_2 - t_1)$$

*Hint:* Check the boundary condition at  $t_2 = t_1$ . Does  $G(2, 1)$  satisfy the given equation when  $t_2 > t_1$ ?

**6.19** The Response of a two-level system driven by an alternating electric field is  $E(t) = E_0 \cos \omega t = E_0/2(e^{i\omega t} + e^{-i\omega t})$ . If two energy levels in an atom are separated by an energy

$$E_2 - E_1 \equiv \hbar\omega_0 \quad \omega_0 > 0$$

with  $\omega_0$  close to the driving frequency  $\omega$ , we may neglect all other energy levels in the atom. We write the time-dependent wave function as

$$\Psi(\mathbf{r}, t) = c_1(t)u_1(\mathbf{r}) + c_2(t)u_2(\mathbf{r}),$$

where  $u_1(\mathbf{r}), u_2(\mathbf{r})$  are the time-independent solutions for the two levels. Thus, if  $H_a$  is the Hamiltonian for the atom without the applied electric field then

$$H_a u_1(\mathbf{r}) = E_1 u_1(\mathbf{r}),$$

$$H_a u_2(\mathbf{r}) = E_2 u_2(\mathbf{r}).$$

However, assuming that the electric field is in the  $z$ -direction, the true Hamiltonian is given by

$$H = H_a - ezE(t).$$

We can define the dipole moment of the atom as

$$\begin{aligned}\mu &= \int u_1^*(\mathbf{r}) e z u_2(\mathbf{r}) d^3 r \\ &= \int u_2^*(\mathbf{r}) e z u_1(\mathbf{r}) d^3 r.\end{aligned}$$

Note

$$\int |u_1(\mathbf{r})|^2 z d^3 r = \int |u_2(\mathbf{r})|^2 z d^3 r = 0.$$

(The two  $u$ 's have opposite symmetries about  $z = 0$ , leading to a nonvanishing dipole moment  $\mu$  coupling the two states.)

a) Use the Schroedinger time-dependent equation to show that

$$\frac{d\varrho_{21}}{dt} = -i\omega_0 \varrho_{21} + \frac{i\mu}{\hbar} E(t)(\varrho_{11} - \varrho_{22}),$$

$$\frac{d\varrho_{22}}{dt} = \frac{-i\mu}{\hbar} E(t)(\varrho_{21} - \varrho_{21}^*),$$

and since  $\varrho_{11} + \varrho_{22} = 1$ ,

$$\frac{d}{dt}(\varrho_{11} - \varrho_{22}) = \frac{2i\mu}{\hbar} E(t)(\varrho_{21} - \varrho_{21}^*).$$

Here,

$$\begin{aligned}\varrho_{11} &= c_1^* c_1 & \varrho_{22} &= c_2^* c_2 \\ \varrho_{21} &= c_1^* c_2 & \varrho_{12} &= c_2^* c_1.\end{aligned}$$

b) In practice, whichever atom is under study collides with other atoms. Here, as one usually does in practice, we are calculating a response which takes cognizance of such collisions. In the absence of the driving field  $E(t)$ , atomic collisions would cause  $\varrho_{11}, \varrho_{22}$  to relax to their equilibrium values  $(\varrho_{11})_0, (\varrho_{22})_0$ . Thus, the final equation in (a) gets an extra term, becoming

$$\frac{d}{dt}(\varrho_{11} - \varrho_{22}) = \frac{2i\mu}{\hbar} E(t)(\varrho_{21} - \varrho_{21}^*) - \frac{(\varrho_{11} - \varrho_{22}) - (\varrho_{11} - \varrho_{22})_0}{\tau}.$$

On the other hand, with collisions,  $\varrho_{21}, \varrho_{12}$  acquire phases specific to each level in each atom. In the absence of the driving field, the presence of phases is considered to cause the averaged values of  $\varrho_{12}, \varrho_{21}$  to vanish with time. Assuming, for simplicity, an equal relaxation time  $\tau$ , for the first equation in (a) we have

$$\frac{d\varrho_{21}}{dt} = -i\omega_0\varrho_{21} + \frac{i\mu}{\hbar}E(t)(\varrho_{11} - \varrho_{22}) - \frac{\varrho_{21}}{\tau}.$$

Show that, on the average, these equations have steady state solutions:

$$\begin{aligned}\varrho_{21}(t) &= \sigma_{21}e^{-i\omega t} \\ \varrho_{12}(t) &= \sigma_{12}e^{i\omega t} = \sigma_{21}^*.\end{aligned}$$

*Hint:* Obtain the equations

$$\begin{aligned}\frac{d\sigma_{21}}{dt} &= i(\omega - \omega_0)\sigma_{21} + \frac{i\mu E_0}{2\hbar}(\varrho_{11} - \varrho_{22}) - \frac{\sigma_{21}}{\tau}, \\ \frac{d}{dt}(\varrho_{11} - \varrho_{22}) &= \frac{i\mu E_0}{\hbar}(\sigma_{21} - \sigma_{21}^*) - \frac{(\varrho_{11} - \varrho_{22}) - (\varrho_{11} - \varrho_{22})_0}{\tau}.\end{aligned}$$

Here, only terms with an  $e^{-i\omega t}$  dependence are retained in the first equation, and no time-dependent exponentials are retained in the second. Such cyclic terms average out to zero. Now we set the left sides equal to zero corresponding to a steady state, and show that they can be solved, i.e., find the constant solutions for complex  $\sigma_{21}$  and  $(\varrho_{11} - \varrho_{22})$  to our inhomogeneous equations.

c) As in Problem 2.10 we define a complex electric susceptibility by

$$P(\omega) = \epsilon_0\chi(\omega)E(\omega)$$

or

$$\begin{aligned}P(t) &= \operatorname{Re}\{\epsilon_0\chi E_0 e^{i\omega t}\} \\ &= E_0(\epsilon_0\chi_1 \cos \omega t - \epsilon_0\chi_2 \sin \omega t) \\ \chi(\omega) &= \chi_1(\omega) + i\chi_2(\omega).\end{aligned}$$

At the same time, we again have

$$P(\omega) = Ne\bar{z}(\omega)$$

with  $N$  atoms per unit volume.

Use part (b) to calculate  $\chi_1(\omega)$  and  $\chi_2(\omega)$ , and compare your results with those obtained in Problem 2.10 where a damped harmonic oscillator was taken as a model for an atom.

**6.20** An important confirmation of quantum theory is derived from measurements of the “Lamb shift”. The shift is in the energy of the  $2s$  level of

hydrogen compared to the  $2p$  level, when both levels are characterized by the same total electron angular momentum. (Its physical basis is the interaction of the electron with the electromagnetic field. Just as an electron interacts with the charged vibrating ions of a host crystal, even if there are no phonons, so can it interact with the EM field even when there are no photons.) The shift itself was measured directly by Lamb and Retherford to be 1057.5 MHZ as compared to the calculated value of 1057.19 MHZ. Estimate the fractional error this discrepancy represents in an actual  $2p \rightarrow 1s$  spectral line.

## Further Reading

H.A. Medicus: *Fifty Years of Matter Waves*, Physics Today (Feb. 1974)

W. Heisenberg: *The Physical Principles of the Quantum Theory* (Dover, New York 1930)

A brief, authoritative statement of the theory.

L.I. Schiff: *Quantum Mechanics* (McGraw-Hill, New York 1968)

This is a classic textbook on quantum mechanics. Many other good texts exist.

P. Carruthers and M.M. Nieto: *Coherent States and the Forced Quantum Oscillator*, Am. J. of Phys. **33**, 537 (1965)

A. Messiah: *Quantum Mechanics* (Dover, Mineola, NY 1999)

G. Rickayzen: *Green's Functions and Condensed Matter* (Academic Press, London and New York 1991)

R.P. Feynman: *The Theory of Fundamental Processes* (W.A. Benjamin, Reading, MA 1961)

S. Fluegge: *Rechenmethoden der Quantentheorie*, 6. Auflage (Springer-Verlag, Berlin, Heidelberg, New York 1999)

Many famous examples in quantum mechanics are worked out here.

Aff. Yariv: *Quantum Electronics*, Third Edition (John Wiley & Sons, New York 1987)

A good place to look for applications of quantum mechanics in a very practical field.

## 7. Nonlinear Waves on Water – Solitons

**Preamble.** There is one kind of wave mentioned in the introduction to Chap. 1 which we have so far pointedly ignored – waves on water. These waves must be among the earliest recorded illustrations of nonlinear behavior, *breaking* noticeably at all but the smallest amplitudes. In previous chapters the equations of motion were linear. Of course, it has been understood that linear equations are idealizations. No matter how small the displacement of a spring, if finite, it must be that its elastic limit is in some fashion exceeded as discussed, for example, by *Erber et al.*, listed under Further Reading. Idealization implies that, although something has been left out for simplicity, the idealized model reflects the essentials of the physical situation. But water waves are different. Results obtained by linearizing, that is, by ignoring the nonlinear parts, are most frequently too far from reality to be useful. In particular, linearization misses a central phenomenon, *solitons*, which are isolated waves or pulses which maintain their identity indefinitely just when we most expect that dispersion effects will lead to their rapid disappearance. Further, water waves act as prototypes for many other nonlinear physical phenomena which, as technology develops are becoming ever more important.

The creation of orderly stable pulses as an effect of nonlinearity has come to science as a surprise. It has also come as a surprise that the more intuitive outcome of nonlinearity, chaos, can be constructively analyzed including transitions from true chaos to structural order. Broadly speaking, Chaps. 2 through 5 deal with phenomena well understood in 1900, Chap. 6 brings one to 1930. Progress in nonlinear dynamics has taken dramatic steps forward since about 1960. Nonlinear waves and solitons are taken up in this chapter. Nonlinearity and chaos are discussed in the essays of Chap. 8.

**Abstract.** We develop the theory of waves on water, starting with linear waves, then giving an account of the effect of nonlinearity, and lastly going on to the existence and description of solitons. In the last section we outline the method of inverse scattering for obtaining multi-soliton solutions. Thus, in Sect. 7.1 we develop the theory of gravity waves on incompressible fluids in the absence of nonlinearity. The picture that emerges is one where in deep channels the elements of a fluid undergo circular motion of diminishing radius as we proceed perpendicularly from the surface into the channel. At the same time, dispersion, the change of the phase velocity of

the waves with wave number, is revealed to be especially important in deep water. Section 7.2 is devoted to exploring the consequences of dispersion, with a discussion contrasting group and phase velocities. Nonlinear behavior is first studied in Sect. 7.3. This section begins with a description of the theory of *characteristics* as applied to the solution of first order, nonlinear partial differential equations (pde). Subsequently, direct application of the method of characteristics to the pdes for nonlinear water waves leads to an understanding of the deformation of the wave forms with time. Section 7.4 introduces solitons. First, it is shown that the combination of dispersion and nonlinearity together leads to the famous Korteweg–de Vries (KdV) partial differential equation. An explicit expression for the single soliton solution of the KdV equation is introduced next, followed by the much more complicated expression for several solitons. In the last section, 7.5, it is shown how the complex solutions of the KdV equation corresponding to many solitons can be generated systematically by the method of inverse scattering. In this method the desired solution of the KdV equation is taken to be the potential in a scattering problem. The potential doing the scattering is reconstructed only after the scattered function, whose time evolution is indirectly controlled by the KdV equation, has been formally determined. It turns out that each soliton can be traced to a single bound eigenstate of the initial scattering potential, that is, of the wave disturbance as it is given at the initial time.

## 7.1 Linear Surface Waves on Water

The waves observed on the surface of a body of water, such as a pond, are very different in nature from sound waves in an elastic solid or in a gas, and the analysis is also different. Elastic waves depend on forces in reaction to compression and rarefaction. A liquid such as water has relatively very small compressibility, and in our analysis we shall leave it out altogether; the fluids supporting the surface waves will be taken as incompressible. It is the combination of adaptability to change of shape which, by definition, fluids possess, and the action of gravity forces which come into play when the fluid is disturbed from its static configuration of minimum energy that gives rise to the surface wave motion.

The analysis shall be restricted to nonviscous (ideal) fluids. Furthermore, eddies or vortices, which in a nonviscous fluid persist without decaying, will be left out. The velocity  $\mathbf{u}(\mathbf{r}, t)$  of any element of the fluid will have zero curl. Consequently, as we know from Chap. 4,  $\mathbf{u}(\mathbf{r}, t)$  can be taken as the gradient of a scalar field  $\phi(\mathbf{r}, t)$  which is the velocity potential:

$$\mathbf{u}(\mathbf{r}, t) = -\nabla\phi(\mathbf{r}, t). \quad (7.1)$$

Also from Chap. 4, we know that if  $\varrho(\mathbf{r}, t)$  is the density of a fluid, the continuity equation will read

$$\varrho \nabla \cdot \mathbf{u} + \partial \varrho / \partial t = 0 ,$$

where  $\varrho \mathbf{u}$  is the current density in mass per unit area – unit time. However, we have just finished saying that  $\varrho$  will be taken as constant,  $\partial \varrho / \partial t = 0$ . The incompressible fluid satisfies

$$\nabla \cdot \mathbf{u} = 0 . \quad (7.2)$$

Evidently (7.1) and (7.2) combine to give *Laplace's equation* (in electrostatics this is the equation satisfied by a potential in a charge-free region):

$$\nabla^2 \phi = 0 . \quad (7.3)$$

The problem we have in mind is based on a body of water confined to a channel or canal of constant width bound by smooth vertical walls and a smooth level bottom. The equilibrium depth of the water is  $h$ . Surface tension, which would give rise to a pressure difference across any curved boundary, will be omitted. Lastly, the amplitudes of our waves will, for now, continue to be small. We are still concerned with only linear waves. (Five effects are being left out! – compressibility, vortices, viscosity, surface tension, and nonlinearity.)

We will be interested in disturbances traveling along the length of the canal, taken to be the  $x$  direction, with  $z$  the direction perpendicular to the surface. Thus, the velocity potential will be a function of  $x$ ,  $z$ , and  $t$ ,  $\phi(x, z, t)$ , signaling “plane” waves, i.e., no variation in the lateral ( $y$ ) direction.

$\phi(x, z, t)$  will be subject to boundary conditions at the bottom  $z = 0$ , and top of the channel,  $z = d$ . It becomes convenient to let  $h(x, t)$  measure the height of the disturbed surface relative to the undisturbed surface at  $z = d$ , where  $d$  is the (average) depth. The condition at  $z = 0$  is simple enough, the vertical velocity  $u_z$  vanishes there:

$$u_z(x, z, t) = -\partial \phi(x, z, t) / \partial z = 0 \quad \text{at } z = 0 . \quad (7.4)$$

On the other hand, at  $z = d$  there will exist an *excess* pressure  $p$ . The hydrostatic pressure  $p_0$ , i.e., the pressure when the water is undisturbed, at a height  $z$  above the bottom of the channel is given by

$$p_0 = p_a + \varrho g(d - z) \quad (7.5a)$$

and, the excess pressure  $p$  is defined as

$$p \equiv p_{\text{tot}} - p_0 , \quad (7.5b)$$

where  $p_a$  is the atmospheric pressure, and  $p_{\text{tot}}$  is the actual pressure. This excess pressure as it exists at the surface  $z = d$ , i.e.,  $p(x, d, t)$ , can be related to the vertical dynamic displacement of the surface  $h(x, t)$  as

$$p(x, d, t) = h(x, t) \varrho g \quad (7.6)$$

where  $g$  is the gravitational constant.

The excess pressure produces acceleration of elements of the fluid. Limiting ourselves in this section to a linear theory, in which the changes in time

in the fluid velocity  $\mathbf{u}(x, z, t)$  due to the corresponding change in location of the element can be neglected, we have

$$\varrho \frac{\partial \mathbf{u}(x, z, t)}{\partial t} = -\nabla p(x, z, t) \quad (7.7)$$

or,

$$\varrho \nabla \partial \phi(x, z, t) / \partial t = -\nabla p(x, z, t).$$

Thus,

$$p = -\varrho \partial \phi / \partial t. \quad (7.8)$$

Combining (7.6) and (7.8) we find

$$[\partial \phi / \partial t]_{z=d} = -h(x, t)g. \quad (7.9)$$

The vertical velocity of the fluid at the surface,  $u_z(x, d, t)$ , may in a linear theory again be equated to  $\partial h(x, t) / \partial t$ , the vertical velocity of the surface. Differentiation of both sides of (7.9) with respect to  $t$ , we find with the help of (7.1) for the second boundary condition

$$\partial^2 \phi / \partial t^2 = -g \partial \phi / \partial z \quad \text{at } z = d. \quad (7.10)$$

The problem thus consists of solving the Laplace equation (7.3)

$$\frac{\partial^2 \phi(x, z, t)}{\partial x^2} + \frac{\partial^2 \phi(x, z, t)}{\partial z^2} = 0 \quad (7.11)$$

subject to the conditions at the bottom and top of the channel, (7.4, 10), respectively. The solution can be readily found by using the method of separation of variables, fully illustrated in Chap. 3 in the analysis of waves on a stretched string. The solution for waves propagating in the forward direction is

$$\phi(x, z, t) = A \cosh kz \cos[kx - \omega(k)t], \quad (7.12a)$$

where

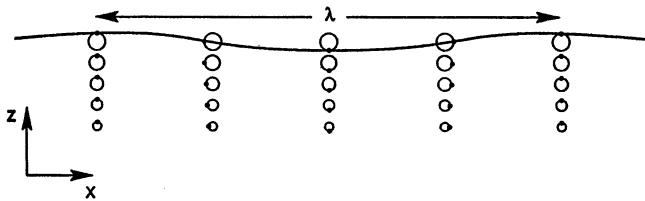
$$\omega^2(k) = gk \tanh kd \quad (7.12b)$$

and  $A$  is the amplitude. The reader can arrive at the details by solving Problem 7.1.

The physical picture corresponding to the solution (7.7) is most easily described for the case of deep water, that is, with the depth  $d$  greater than one wavelength,  $2\pi/k$ . Here we can replace  $\sinh kz, \cosh kz$  by  $e^{kz}/2$  near the surface at  $z = d$ . The two velocity components  $u_x, u_z$  become

$$u_x = -\frac{\partial \phi}{\partial x} = \frac{kA}{2} e^{kz} \sin(kx - \omega t), \quad (7.13a)$$

$$u_z = -\frac{\partial \phi}{\partial z} = \frac{kA}{2} e^{kz} \cos(kx - \omega t). \quad (7.13b)$$



**Fig. 7.1.** Fluid elements of deep water undergoing circular motion, with phases depending on  $x$ , leads to the familiar surface waves on water. This figure is given by J. Lighthill in his *Waves in Fluids*

An element of fluid is seen to be undergoing circular motion (clockwise since  $u_z$  is leading) of fixed radius, with only the phase depending on  $x$ . With increasing depth ( $z$  becoming smaller than  $z = d$ ) the radius diminishes. Figure 7.1 gives a complete picture of how the surface waves work.

## 7.2 Dispersion. Group Velocity

As long as the differential equation governing a system is linear, one can obtain solutions as four dimensional (space and time) Fourier representations with harmonic functions:

$$e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)} \quad \text{or} \quad \sin(\mathbf{k} \cdot \mathbf{r} - \omega t), \cos(\mathbf{k} \cdot \mathbf{r} - \omega t).$$

The ensuing equation connecting  $\omega$  and  $k$  is called the *dispersion relation* for the system. The simplest form of the dispersion relation is a linear one. Systems with this relationship are said to be *nondispersive*. The first systems we studied, idealized waves on a stretched string in Chap. 3, electromagnetic waves in a vacuum in Chap. 4, were nondispersive. However, we have also seen three instances of *dispersive* systems, EM waves propagating in a medium, as in Chap. 5, the matter waves of Chap. 6, and the water waves of this chapter, where  $\omega$  is related to  $k$  by means of (7.12b).

As we found in Chap. 3, whenever we have a harmonic solution, the velocity of the wave, for example, of the antinode, is given by

$$s_{\text{ph}} = (\omega/k). \tag{7.14}$$

The velocity defined by (7.14) is called the *phase velocity*, and for nondispersive systems it is a constant independent of  $k$  or  $\omega$ . The fact is that in physics we are seldom interested in sending harmonic disturbances but rather in sending distinctive physical signals. As we saw in Sect. 6.3, it is convenient to represent such signals by their Fourier integrals (wave packets). It is clear that since with linear dispersion relations all the wave components travel with the same (phase) velocity, the signal maintains its original shape; it does not “disperse”, in the commonly accepted meaning of that word. (The Appleton–Century dictionary defines the verb “disperse” when intransitive

as: “to separate or move apart in various directions; become dissipated or dispelled”.) On the other hand, in those situations where the dispersion relation is nonlinear, different harmonic components will travel at different (phase) velocities, and the signal does disperse; the energy associated with the signal disperses as well; the systems are dispersive!

It now becomes natural to ask if one can find a unique velocity for the signal, or velocity of the wave packet while it has not yet completely dispersed. In one dimension a suitable wave packet can be represented by the Fourier integral

$$h(x, t) = 1/\sqrt{2\pi} \int_{-\infty}^{\infty} A(k) \exp[i[kx - \omega(k)t]] dk . \quad (7.15)$$

We can argue that the principal contribution for any value of  $x$  and  $t$  to this integral comes from the neighborhood of a value  $k_0(x, t)$  found by letting

$$\frac{d}{dk} [kx - \omega(k)t] \Big|_{k=k_0} = 0 , \quad (7.16a)$$

or,

$$\frac{d\omega(k)}{dk} \Big|_{k=k_0(x, t)} = \frac{x}{t} . \quad (7.16b)$$

Uniquely in the neighborhood of this value of  $k_0$  do the phases add constructively (method of stationary phase). Equation (7.16b) is of direct value whenever the signals of interest can be described by rather tight packets centered about some fixed wave vector  $k_0$ . One can then turn (7.16b) around and argue that the location of the signal  $x$  is related to the time  $t$  through

$$\frac{x}{t} \equiv s_g = \frac{d\omega(k)}{dk} \Big|_{k=k_0} , \quad (7.17)$$

where the *group velocity*  $s_g$  is truly the velocity of the signal.

Consider now the situation when a stone is thrown into a “deep” pond. The dispersion relation for water waves in general is

$$\omega(k) = [gk \tanh kd]^{1/2} . \quad (7.12b)$$

By deep we mean  $kd \gg 1$ , i.e., the depth of the water corresponds to many wavelengths. Here,

$$\begin{aligned} \tanh kd &\equiv \frac{(e^{kd} - e^{-kd})}{(e^{kd} + e^{-kd})} \approx 1 \\ \omega(k) &\approx (gk)^{1/2} \\ s_g(k) &= \partial\omega/\partial k = \frac{1}{2}(g/k)^{1/2} \\ s_{ph}(k) &= \omega/k = (g/k)^{1/2} . \end{aligned} \quad (7.18a) \quad (7.18b)$$

Thus, if we observe waves with wavelengths distributed about some central value, we shall find that they do not spread out at their phase velocities at all, but at half of those velocities! Clearly, the distinction between  $s_{\text{ph}}$  and  $s_g$  is important.

The concept of group velocity is helpful in interpreting the propagation of wave packets in quantum mechanics, as in Sect. 6.3 (Problem 7.2). In advanced studies of classical wave phenomena the group velocity indeed plays an important role (refer to the texts for further reading listed at the end of the chapter).

By contrast with (7.18) for shallow water (long waves,  $kd \ll 1$ ), we have

$$s_g \equiv s_0 = \partial\omega/\partial k = (gd)^{1/2}. \quad (7.19)$$

The propagation is without dispersion (Problem 7.3).

For somewhat shorter waves (“fairly long waves”) we expand (7.12b) in  $kd$ , stopping after two terms (Problem 7.4) and find

$$s_{\text{ph}} = \omega/k = s_0 \left(1 - \frac{1}{6}k^2 d^2\right). \quad (7.20)$$

## 7.3 Nonlinear Waves

The acceleration  $\mathbf{a}(\mathbf{r}, t)$  of an element of fluid is given in Cartesian coordinates in fact by three equations; for the  $x$  component we have

$$a_x = \frac{\partial u_x}{\partial t} + u_x \frac{\partial u_x}{\partial x}. \quad (7.21)$$

The first term on the right gives the rate of change with time at a fixed point, while the second term takes account of the element’s changing location. It is this second term which has been omitted in the previous two sections so as to have a linear formulation. The purpose of this section is to explore the consequences of the nonlinear term. To simplify we omit here several other features which were considered previously, namely, the two dimensional character of waves in open channels, and dispersion.

We restrict ourselves to shallow waves where, as we have seen, dispersion effects become negligible. When analyzing deep water waves we found that elements near the surface undergo circular motion. Suppose, on the contrary, that the water is shallow compared to the wavelength  $2\pi/k$ . Referring to (7.12a) for the velocity potential  $\phi(x, z, t)$ , we see that the two velocity components  $u_x, u_z$ , both shown in (7.13a) and (7.13b) as depending on  $e^{kz}$ , are actually given by  $\cosh kz$  and  $\sinh kz$ , respectively. This means that for shallow waves,  $kd \gg 1$ ,  $u_x$  much exceeds  $u_z$  near the surface at  $z = d$ . As a consequence, we can treat the propagation as one dimensional, and can find a solution directly without resorting to a velocity potential. The cross-sectional area of the channel  $A(p)$ , however, must be allowed to vary with the excess pressure as before. We had

$$p(x, t) = \varrho g h(x, t), \quad (7.6)$$

and,

$$A(x, t) = b[d + h(x, t)] = bd + p(x, t)b/\varrho g \quad (7.22a)$$

or

$$A(p) = bd + pb/\varrho g, \quad (7.22b)$$

where  $d$  is the equilibrium depth of the channel, and  $b$  its constant width.

As in (7.7) the gradient of the excess pressure gives rise to acceleration of elements of the fluid. In view of (7.21) we have, in one dimension,

$$\varrho(\partial u/\partial t + u\partial u/\partial x) = -\partial p/\partial x \quad (7.23)$$

where  $\varrho$  is the constant density of the fluid. The other key equation is again the equation of continuity, which for an incompressible fluid in a channel of varying cross section  $A$  is given by

$$\varrho\partial A/\partial t + \varrho\partial(Au)/\partial x = 0. \quad (7.24)$$

Since the time of the French mathematician Augustin Cauchy (1789–1857), it has been known that certain initial value problems involving partial differential equations, even nonlinear equations, can be solved by first finding the so-called *characteristics*. For our purposes it will suffice to illustrate the method for an initial value problem where the unknown function  $w(x, t)$  satisfies a partial differential equation (pde) of the form

$$a(x, t, w)\frac{\partial w(x, t)}{\partial x} + b(x, t, w)\frac{\partial w(x, t)}{\partial t} = 0, \quad (7.25)$$

and  $w(x, 0)$  is prescribed. Since  $a$  and  $b$  may depend on  $w$ , this equation is nonlinear. A characteristic for the pde (7.25) is any function  $x(t)$  obtained by solving the ordinary differential equation

$$\frac{dx}{dt} = \frac{a(x, t, w)}{b(x, t, w)}, \quad (7.26)$$

where  $w$  is to be treated as a constant parameter [ $x(0)$  must be within the limits for which  $w(x, 0)$  is prescribed]. We now note that on the characteristic

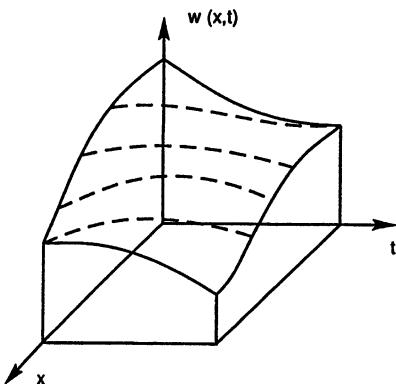
$$dw = \frac{\partial w(x, t)}{\partial x} dx + \frac{\partial w(x, t)}{\partial t} dt,$$

with

$$\frac{dx}{a(x, t, w)} = \frac{dt}{b(x, t, w)} \equiv dl,$$

$$dw = \left[ \frac{\partial w(x, t)}{\partial x} a(x, t, w) + \frac{\partial w(x, t)}{\partial t} b(x, t, w) \right] dl.$$

Hence, if  $w(x, t)$  is a solution of the original equation (7.25),  $dw$  is zero, that is,  $w$  is a constant on the characteristic curve  $x(t)$ .



**Fig. 7.2.** The solution surface  $w(x,t)$ . The characteristics, paths of walkers on the surface who do not wish to “climb”, are shown *dashed*. Strictly, functions  $x(t)$  are the projections of the characteristics on the  $xt$  plane, rather than the actual characteristics

Thus, for a homogeneous pde such as (7.25) the characteristics can be defined as paths of constant  $w$  (Fig. 7.2). Moreover, if the characteristics do not cross, then, in general, if we can solve the ordinary differential equation (7.26), we have solved the problem. For if we wish to know  $w(x_0, t_0)$  at some point  $x_0, t_0$ , we simply trace the characteristic passing through this point back to the point  $x(0)$  at  $t = 0$ , i.e.,  $w(x_0, t_0) = w[x(0), 0]$ . All this may remind the reader of Chap. 3 in which we found that a solution of the homogeneous wave equation (3.3) with zero initial velocity was any function of  $x \pm st$ , i.e., was a constant along these curves;  $x = \pm st$  are characteristics for (3.3), see Problem (7.7).

We should like now to combine (7.22, 23, 24) into first order pdes, each in terms of one unknown function which will be solvable by the method of characteristics. We proceed by defining certain auxiliary functions  $s(p)$  and  $P(p)$ , both having dimensions of velocity  $u$ . First, we notice that if we linearize (7.24) we expect to obtain the same phase velocity as we found for long waves in Sect. 7.2 (Problem 7.8):

$$s_0 = (gd)^{1/2}. \quad (7.19)$$

The convenient generalization of  $s_0$  turns out to be

$$s(p)^{-2} \equiv [\varrho/A(p)]dA(p)/dp. \quad (7.27a)$$

From (7.19),

$$s^{-2} = b/Ag \approx s_0^{-2}. \quad (7.27b)$$

The physical significance of  $s$  will become apparent shortly. Second, we define an integral  $P(p)$  by

$$P(p) \equiv 1/\varrho \int_0^p \frac{1}{s(p')} dp' \quad (7.28)$$

where  $s(p)$  is defined by (7.22) and (7.27).

Going back now to the equation of motion (7.23) we first notice that from (7.28)

$$\varrho \frac{\partial P}{\partial x} = \frac{1}{s(p)} \frac{\partial p}{\partial x}, \quad \varrho \frac{\partial P}{\partial t} = \frac{1}{s(p)} \frac{\partial p}{\partial t}$$

giving, after division by  $\varrho$

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + s \frac{\partial P}{\partial x} = 0. \quad (7.29a)$$

Dividing the equation of continuity (7.24) by  $\varrho A/s$  we obtain

$$\frac{s}{A} \frac{\partial A}{\partial t} + s \frac{\partial u}{\partial x} + \frac{s}{A} u \frac{\partial A}{\partial x} = 0.$$

Noting also that

$$\varrho \frac{\partial P}{\partial t} = \frac{1}{s(p)} \frac{\partial p}{\partial t}, \quad \frac{\partial A}{\partial t} = \frac{dA}{dp} \frac{\partial p}{\partial t}, \quad \text{and} \quad \frac{\partial A}{\partial x} = \frac{dA}{dp} \frac{\partial p}{\partial x},$$

and making use of the definition of  $s^{-2}$  in (7.27), we can now rewrite the equation of continuity as

$$\frac{\partial P}{\partial t} + s \frac{\partial u}{\partial x} + u \frac{\partial P}{\partial x} = 0. \quad (7.29b)$$

Adding (7.29a) and (7.29b) we obtain,

$$\frac{\partial(u + P)}{\partial t} + (u + s) \frac{\partial(u + P)}{\partial x} = 0, \quad (7.30a)$$

and subtracting,

$$\frac{\partial(u - P)}{\partial t} + (u - s) \frac{\partial(u - P)}{\partial x} = 0. \quad (7.30b)$$

Equations (7.30a,b) constitute the desired nonlinear first order pdes that govern the wave motion.

In the remainder of this section we shall use our knowledge of the theory of characteristics to obtain a general insight into the propagation of nonlinear waves without solving any specific problem analytically. Let us call the characteristics associated with (7.30a)  $C^+$ . On these,  $u + P$  is constant, and from (7.26)

$$\frac{dx}{dt} = u(x, t) + s(x, t) \quad \text{on} \quad C^+. \quad (7.31a)$$

For (7.30b) we have, similarly,

$$\frac{dx}{dt} = u(x, t) - s(x, t) \quad \text{on} \quad C^-, \quad (7.31b)$$

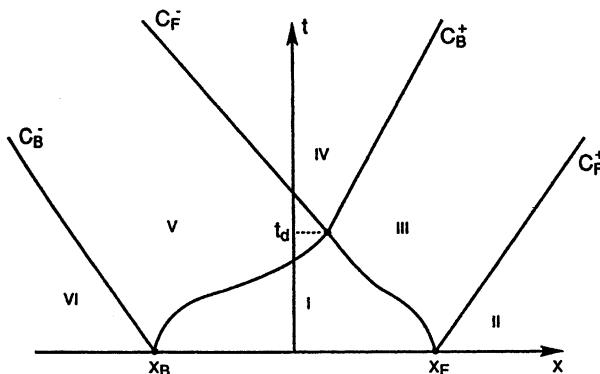
with  $u - P$  constant.

Consider the initial value problem where at  $t = 0$  all the fluid in the channel is at rest except in the interval  $x_F > x > x_B$ :

$$u(x, 0), p(x, 0), P(p = 0) = 0 \quad x > x_F, \quad x < x_B \quad (7.32a)$$

$$u(x, 0), p(x, 0), P(p = 0) \neq 0 \quad x_F \geq x \geq x_B. \quad (7.32b)$$

The characteristics  $C^+$ ,  $C^-$  originating at  $x_F$  and  $x_B$ , respectively, divide the  $xt$  plane into six zones, Fig. 7.3. Regions II, IV, and VI in the figure do not have any associated disturbance;  $u(x, t)$  vanishes. Region II is free because the most forward part of the disturbance is bound by  $C_F^+$ ;  $u + P$ , which is constant on  $C^+$ , vanishes on all more forward  $C^+$  characteristics in view of (7.32a). The disturbance has not yet reached the space-time points in II. The same holds in VI, as one can show by reversing all directions. Further, the complex disturbance in I separates or disentangles into two disturbances moving in opposite directions, Regions III and V, gradually leaving the original region  $x_B > x > x_F$ , and its subsequent extension IV clear of the motion, Problem 7.9. The disturbances in Regions III and V are referred to as *simple waves*; it is to these that we turn next.



**Fig. 7.3.** Characteristics originating at  $x_F$  and  $x_B$ , respectively, divide the  $xt$  plane into six zones. Water in Zones II, IV and VI is undisturbed. Zone I becomes empty at time  $t_d$ , when the original disturbance entirely disentangles into the “simple waves” found in III and V. Here characteristics are straight lines, but, in general, not parallel. This figure is adapted from J. Lighthill’s *Waves in Fluids*

We can imagine Region III covered by  $C^-$  characteristics originating at  $x > x_F$ , where  $u$  and  $P$  separately vanish. Thus, since on  $C^-$  the function  $u - P$  is constant, we have

$$u(x, t) = P(x, t) \quad \text{in III} \quad (7.33a)$$

so that

$$u(x, t) + P(x, t) = 2u(x, t) \quad \text{along } C^+ \text{ in III.} \quad (7.33b)$$

From the respective definitions (7.27, 28) of  $s$  and  $P$ , and (7.22b) giving  $A(p)$ , we see again that  $s$  and  $P$  are functions of  $p$  only. Thus, viewing (7.33b) we

conclude that  $u(x, t)$ ,  $P(x, t)$ ,  $s(p)$ ,  $p$  are all constants along any characteristic  $C^+$  in III; by (7.31a) these characteristics become straight lines. It is this latter fact that prompts us to refer to the disturbances in III as “simple waves”. Similarly, in Zone V the  $C^-$  characteristics will also be straight lines, and we have simple waves there as well.

However, the characteristics in III and V are not parallel as they are for linear waves (Problem 7.9). Using (7.28) for  $P(p)$  with (7.27b) for  $s$ , we have for long waves in channels (with small modifications the theory holds for many other phenomena, e.g., blood flowing in distendable arteries)

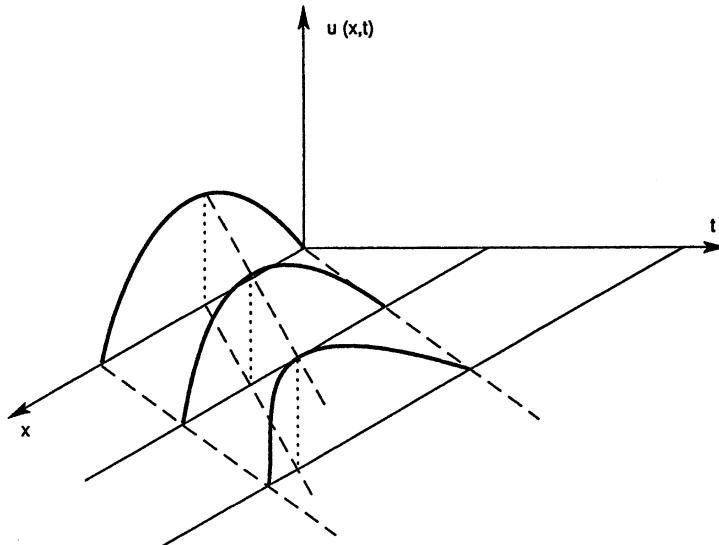
$$P(p) = 1/\varrho \int_0^p \frac{dp}{s(p)} = 1/\varrho \sqrt{\frac{b}{g}} \int_0^p \frac{dp}{A^{1/2}} .$$

With (7.22b) for  $A(p)$  this equation becomes

$$P(p) = \sqrt{g/b} \int_{A(0)}^{A(p)} A^{-1/2} dA ,$$

or

$$P(p) = 2[s(p) - s_0] . \quad (7.34a)$$



**Fig. 7.4.** Progression of “simple” nonlinear waves. The characteristics, shown dotted, are not parallel, leading to the development of a sharper leading edge. Eventually the wave will break

Combined with (7.33a), (7.34a) gives

$$s = s_0 + u/2 \quad \text{or} \quad u + s - s_0 = \frac{3}{2}u . \quad (7.34b)$$

As a result, the slope of each characteristic  $C^+$ , (7.31a) becomes in III

$$\frac{dx}{dt} = s_0 + \frac{3}{2}u(x, t) \quad \text{in III} . \quad (7.35)$$

The larger the  $u$  the more rapid the progression of any point (particular value of  $u$ ) on the wave. The consequence of the differing slopes of characteristics is that nonlinear waves distort as they progress, as illustrated in Fig. 7.4. Eventually, Fig. 7.5b, the wave breaks.

## 7.4 Solitons

The physics of waves on water that we have been learning in the previous sections of this chapter is illustrated in Fig. 7.5a,b, and can be summarized analytically in the so-called *Korteweg-de Vries equation*. According to (7.31a) we can interpret  $(u + s)$  as a “signal” velocity, and  $s$  as the velocity of the signal relative to the motion of the fluid, as given by  $u$ . Let us call the velocity with which the wave is deforming  $v(x, t)$ :

$$v(x, t) \equiv u(x, t) + s(x, t) - s_0 = \frac{3}{2}u(x, t) , \quad (7.36)$$

where  $s_0$  is the linear velocity, i.e., the limiting value of  $s$  as  $u$  goes to zero, (7.34b). At the same time we obtained in the last section that  $u + p = 2u$ , so that

$$u(x, t) + P(x, t) = \frac{4}{3}v(x, t) .$$

As a result (7.30a) assumes the very convenient form

$$\frac{\partial v(x, t)}{\partial t} + [v(x, t) + s_0] \frac{\partial v(x, t)}{\partial x} = 0 . \quad (7.37)$$

In Sect. 7.2 we saw that for waves in water which is not shallow compared to the wavelength, dispersion occurs, with a phase velocity

$$s = s_0 \left( 1 - \frac{1}{6}k^2 d^2 \right) . \quad (7.14)$$

This dispersion can be included by adding into (7.37) the linear term  $\sigma \partial^3 v(x, t)/\partial x^3$ , yielding

$$\frac{\partial v}{\partial t} + (s_0 + v) \frac{\partial v}{\partial x} + \sigma \frac{\partial^3 v}{\partial x^3} = 0 , \quad (7.38a)$$

$$\sigma = (1/6)s_0 d^2 . \quad (7.38b)$$

Equations (7.38a,b) are the Korteweg–de Vries equation, which we will be studying for the remainder of this chapter. However, it will be worthwhile

to put (7.38a,b) into a slightly different form to conform with the usage of applied mathematicians, thereby facilitating any interested reader's access to current research. Three minor mathematical manipulations are needed. The first is to transform to a new coordinate system, itself moving at speed  $s_0$ . Thus

$$x_{\text{new}} = x_{\text{old}} - s_0 t_{\text{old}} \quad (7.39a)$$

$$t_{\text{new}} = t_{\text{old}} . \quad (7.39b)$$

The second is to replace  $v$ :

$$v_{\text{new}} = -v_{\text{old}} . \quad (7.39c)$$

The third and last will be to make a change of scale in  $x$  resulting in  $\sigma$ 's replacement by unity in (7.38a). Noting that

$$\frac{\partial}{\partial t_{\text{old}}} = \frac{\partial}{\partial t_{\text{new}}} \frac{\partial t_{\text{new}}}{\partial t_{\text{old}}} + \frac{\partial}{\partial x_{\text{new}}} \frac{\partial x_{\text{new}}}{\partial t_{\text{old}}} = \frac{\partial}{\partial t} - s_0 \frac{\partial}{\partial x}$$

$$\frac{\partial}{\partial x_{\text{old}}} = \frac{\partial}{\partial t_{\text{new}}} \frac{\partial t_{\text{new}}}{\partial x_{\text{old}}} + \frac{\partial}{\partial x_{\text{new}}} \frac{\partial x_{\text{new}}}{\partial x_{\text{old}}} = \frac{\partial}{\partial x} ,$$

where we drop subscripts on the new quantities on the right and below, (7.38a) becomes the Korteweg–de Vries equation in its most usual form:

$$\frac{\partial v(x, t)}{\partial t} - 6v(x, t) \frac{\partial v(x, t)}{\partial x} + \frac{\partial^3 v(x, t)}{\partial x^3} = 0 .$$

(7.40)

Fairly long surface waves on water actually do satisfy (7.40).

At first it might be supposed that nothing of new significance can be learned from this equation; the effects of dispersion and nonlinearity have already been explored. Indeed, given some initial disturbance  $v(x, 0)$ , (7.40) has, in general, solutions  $v(x, t)$  which die out in time in a dispersive manner. However, it turns out that it is also possible to create an excitation which remains stable, one that does not dissipate its energy with time – the soliton.

The occurrence of a soliton was first recorded by the British civil engineer J. Scott Russell in 1834 who wrote the now famous description: "I was observing the motion of a boat which was rapidly drawn along a narrow channel by a pair of horses, when the boat suddenly stopped – not so the mass of water in the channel which it had put in motion; it accumulated round the prow of the vessel in a state of violent agitation, then suddenly leaving it behind, rolled forward with great velocity, assuming the form of a large solitary elevation, a rounded, smooth and well-defined heap of water, which continued its course along the channel apparently without change of form or diminution of speed. I followed it on horseback, and overtook it still rolling on at a rate of some eight or nine miles an hour, preserving its original figure some thirty feet long and a foot to a foot and a half in height. Its height

gradually diminished, and after a chase of one or two miles I lost it in the windings of the channel."

Solitons are not confined to shallow water waves. All physical phenomena go nonlinear at sufficiently high amplitudes. Theoretical physicists, perhaps, too long tended to avoid the problems raised by nonlinearity. The techniques we have learned in previous chapters for linear systems, Fourier analysis, or mode superposition, do not work for nonlinear problems and, as is becoming increasingly clear, nonlinearity often leads to chaotic behavior. Yet an orderly outcome, the soliton, is also possible. There are ion-plasma-wave solitons, high-intensity shock solitons, nerve conduction solitons. Solitons have now been observed in virtually every branch of physics. In nuclear physics inverse methods have been used to derive soliton equations and solutions from the Schroedinger eigenvalue problem. In condensed matter physics so-called topological solitons whose passage brings a change of state are common; in superconductivity, vortices carrying unit magnetic flux  $h/2e$  are solitons, interfaces between different structural phases in crystals, domain walls in magnetism, some dislocations are all solitons. At the time of this writing a multi-billion dollar project for an optical fiber cable across the Pacific to carry signals as optical fiber solitons is pending.

The stability of solitons can, to some extent, be understood as due to a combination of dispersive and nonlinear effects. In Sect. 7.2 we explored the effect of dispersion. We found that for water waves, long wave components travel faster than short wave components, leading to a gradual dissipation of a pulse, as shown in Fig. 7.5a. At the same time, nonlinear effects were shown in Sect. 7.3 to lead to a concentration of the pulse on the leading side of a wave, as shown in Fig. 7.5b. The existence of a stable solitary wave (soliton), indicated by the pulse in Fig. 7.5c, may be understood as resulting from an opportune balance of these two effects.

An analytic solution of the KdV equation (Problem 7.10) corresponding to a single soliton can be conveniently written as

$$v(x, t) = -\alpha^2/2 \operatorname{sech}^2 [\alpha/2(x - \alpha^2 t - x_1)] . \quad (7.41)$$

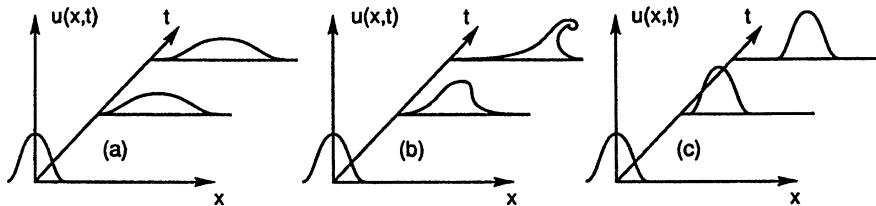
Alternatively, it is helpful to define an  $f(x, t)$  as

$$f(x, t) \equiv e^{-\alpha(x-x_1)+\alpha^3 t} \quad (7.42a)$$

in which case,

$$v(x, t) = \frac{-2\alpha^2 f(x, t)}{[1 + f(x, t)]^2} . \quad (7.42b)$$

The maximum absolute value of  $v(x, t)$ , the amplitude, is given by  $\alpha^2/2$ , occurring at  $f = 1$ , that is, at  $x = \alpha^2 t + x_1$ , where  $x_1$  is the phase shift. The velocity of the soliton is given by  $\alpha^2$ , that is, by twice its amplitude, a very nonlinear result. (It should be remembered, however, that to find the actual velocity of a pulse such as shown in Fig. 7.5c we must add in  $s_0$ .) Initial



**Fig. 7.5.** (a) Water waves in anything but very shallow channels disperse. (b) Nonlinear effects produce a sharpening of the leading edge of a wave, ultimately leading to breakers. (c) When the dispersion shown in (a) is just balanced by the concentration in (b) the stable soliton results

conditions determine both  $\alpha$  and  $x_1$ , the two unknown parameters, as will be shown in the next section.

So far we have emphasized the stability of a single soliton; dispersion effects are baffled. Again, the disturbance is nonlinear, the amplitude determines the velocity. However, we remember from Chap. 2 that the hallmark of linearity is superposition: If we generate one homogeneous (undriven) solution, and then a second solution, the sum of the two solutions likewise becomes a solution. The two solutions do not affect each other. Looking at (7.40), the KdV equation, such superposition clearly can not be expected. Now, if we have two solitons of different amplitudes, they will also have different velocities, and we can easily arrange for the larger soliton to catch up and bump into the smaller one. The very remarkable thing is that both solitons will survive such a collision unscathed, the larger passing the smaller; the only effect is a change of phase for each soliton. The soliton, indeed, appears in some ways as the KdV counterpart of a linear disturbance. How can we prove analytically the stability to collisions, which actually holds independent of the number of solitons present? The analytic solution with two solitons in terms of parameters  $\alpha_1, \alpha_2$  defining  $f_1, f_2$ , respectively, is given by

$$v(x, t) = \frac{\alpha_1^2 f_1 + \alpha_2^2 f_2 + 2(\alpha_2 - \alpha_1)f_1 f_2 + \Lambda (\alpha_2^2 f_1^2 f_2 + \alpha_1^2 f_1 f_2^2)}{(1 + f_1 + f_2 + \Lambda f_1 f_2)^2} \quad (7.43a)$$

$$\Lambda \equiv (\alpha_2 - \alpha_1)^2 / (\alpha_2 + \alpha_1)^2, \quad (7.43b)$$

where  $f_1$  and  $f_2$  are also given by (7.42a), each  $f$  having its own arbitrary parameters  $\alpha_1, x_1$ , and  $\alpha_2, x_2$ , respectively. The demonstration that  $v(x, t)$  as given in (7.43) indeed corresponds to nondeforming collisions of two solitons is given in Problem 7.11. The point to be noted here is that (7.43), with all its complexity is an exact solution of the KdV equation (7.40). How does this expression look for  $N$  solitons? How can such complicated solutions be found? Can we obtain a deeper understanding? It is this understanding which forms the subject of the next, last, section of this chapter.

## 7.5 Inverse Scattering

Inverse scattering is a method for solving certain nonlinear pdes, for example, the Korteweg–de Vries equation. The desired solution of the pde is taken as the potential in an artificially constructed scattering problem, which is solved as a first step. The time evolution of the scattering solution follows that of the potential, thus being determined indirectly by the original pde equation. In the second step the potential, i.e., the solution of the pde equation, is obtained by a process of reconstruction from the scattering solution. That such a procedure could be used was discovered by Gardener, Greene, Kruskal and Miura in 1967 at the Plasma Physics Laboratory of Princeton University in their work with the KdV equation. This development came in the wake of the first suggestion of the modern ideas for finding soliton solutions in the 1956 paper by the american mathematicians Irvin Kay and Harry Moses, and by the numerical discovery of stable solutions (solitons) of the KdV equation by Zabusky and Kruskal in the course of computer calculations. The point is that one can find the time dependence of the scattering solution without solving the pde itself. Although roundabout, no other method is in general known for finding solutions to these nonlinear pdes. There is, of course, no general method for solving an arbitrary nonlinear pde. However, the equations that can be solved with the method appear rather frequently in physics. Solitons, as we have seen, occur in many physical phenomena. The discussion which follows will be restricted to the Korteweg–de Vries pde.

In the method, the scattering process is represented by the following equation:

$$\frac{\partial^2 \phi(x, \tau; t)}{\partial x^2} - v(x, t)\phi(x, \tau; t) = \frac{\partial^2 \phi(x, \tau; t)}{\partial \tau^2}, \quad (7.44)$$

where  $v(x, t)$  is the desired solution of the KdV equation.  $\phi(x, \tau; t)$  is referred to as the “wave function” of the “scattered particle” and, in further analogy to the Schroedinger equation, we refer to  $v(x, t)$  as the scattering potential. In addition, we shall refer to  $\tau$  as “time”, while regarding the physical time  $t$  over which  $v(x, t)$  evolves as a constant parameter during the scattering process.

Equation (7.44) can be solved by the usual method of separation of variables. If we substitute as a trial function

$$\text{trial function} = u[x, \omega(t); t]e^{-i\omega(t)\tau}, \quad (7.45)$$

$u$  independent of  $\tau$ , we obtain the eigenvalue equation

$$\frac{\partial^2 u[x, \omega(t); t]}{\partial x^2} + [\omega^2(t) - v(x, t)] u[x, \omega(t); t] = 0. \quad (7.46)$$

Motivated by the stability of the soliton one is led to inquire into conditions on  $v(x, t)$  which would make the eigenvalues of (7.46), i.e.,  $\omega^2(t)$ ,

independent of  $t$ . One then discovers the fact, essential to the method, that a sufficient condition for  $\omega^2(t)$  for bound states to be time independent is for  $v(x, t)$  to satisfy the KdV equation, see Problem 7.15. As it turns out, with KdV one can arrange for the eigenvalues of the free states of (7.46) to be constant as well. Thus  $\omega(t)$  will henceforth be written as time independent, anticipating an actual proof.

The particular inverse scattering method to be outlined in detail is that of G.N. Balanis, following upon a suggestion of the same approach in a 1960 paper by Irvin Kay. Here the incident particle is described by a delta function pulse,  $\delta(x + \tau)$ , at every fixed value of the parameter  $t$ . The particle is moving to the left at unit velocity [corresponding to (7.44), where the velocity has also been set equal to unity]. Accordingly, we say that for all  $t$  in the asymptotic limit,

$$\lim_{x \rightarrow \infty} \phi(x, \tau; t) \equiv \phi_\infty(x, \tau; t) = \delta(x + \tau) + B(x - \tau; t) , \quad (7.47)$$

where  $B(x - \tau; t)$  is the reflected signal, the wave scattered by  $v(x, t)$ ; (7.47) is the scattering solution of (7.44) for  $x \rightarrow \infty$ .

It is convenient to classify the eigenfunctions of (7.46) according to their *asymptotic* forms. Scattering implies that  $v(x, t)$  only extends over a finite region  $L > x > -L$ ; outside this region  $v(x, t)$  equals zero. The asymptotic forms of  $u$  and of  $\phi$ , ( $\phi_\infty$ ), are their forms outside the region. We shall have to consider, as already implied, both bound and free eigenstates; the latter are referred to as scattering states.

The bound states decay exponentially in the asymptotic region:

$$u(x, \omega; t) \equiv u_n(x; t) \rightarrow \gamma_n(t) e^{-K_n x} \quad \text{as } x \rightarrow \infty , \quad (7.48a)$$

$$\omega = \mp i K_n , \quad K_n > 0 , \quad (7.48b)$$

i.e., the “total energy”  $\omega^2$  is negative. Either sign is possible for  $\omega$ , see below. The “constants”  $\gamma_n(t)$  are determined by the normalization of the exact eigenfunctions  $u_n(x; t)$ . These states can be taken as real, and are orthonormal:

$$\int_{-\infty}^{\infty} u_n(x; t) u_m(x; t) dx = \delta_{nm} . \quad (7.48c)$$

For the free eigenstates we expect that

$$\text{the state} \rightarrow 1/\sqrt{2\pi} e^{-ikx} \quad \text{as } x \rightarrow \infty .$$

Actually, every ingoing wave will also have an associated scattered wave heading outward from the scattering potential (Problem 6.11), and in the region  $x \rightarrow -\infty$  there will appear a transmitted wave  $e^{ikx}$ . Thus,

$$\begin{aligned} u(x, \omega; t) \equiv u(x, k; t) &\longrightarrow \frac{1}{\sqrt{2\pi}} [e^{-ikx} + \beta(k; t)e^{ikx}] \quad x \rightarrow \infty, k > 0 \\ &\longrightarrow \frac{1}{\sqrt{2\pi}} \alpha(k; t)e^{-ikx} \quad x \rightarrow \infty, \quad k < 0 \end{aligned} \quad (7.49a)$$

$$\begin{aligned} &\longrightarrow \frac{1}{\sqrt{2\pi}} \alpha(k; t)e^{-ikx} \quad x \rightarrow -\infty, \quad k > 0 \\ &\longrightarrow \frac{1}{\sqrt{2\pi}} [e^{-ikx} + \beta(k; t)e^{ikx}] \end{aligned}$$

$$x \rightarrow -\infty, k < 0 \quad (7.49b)$$

$$\omega = \pm k. \quad (7.49c)$$

Here  $\omega$ , unlike the case for the bound states, forms a continuum. The most general solution will contain combinations of positive and negative frequency exponentials for any given value of  $k$  or  $K_n$ . As mentioned we shall be able to find solutions  $\phi(x, \tau; t)$  using *scattering states*  $u(x, k; t)$  with  $k$  independent of  $t$ . The scattering states thus have as their asymptotic form plane waves with, by (7.49c), again unit phase velocities ( $\pm\omega/k$ ). The orthonormality condition may, if one wishes to return to the rigorous formulation of (1.32b), be written as

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} u^*(x, k; t) u(x, k'; t) e^{-x^2/4n} dx = \lim_{n \rightarrow \infty} g_n(k - k'), \quad (7.49d)$$

where the contribution to the integral from the scattering region  $L > x > -L$  can be taken as going to zero. The parameters  $\alpha, \beta$  in (7.49a,b) are called the scattering coefficients. From Chap. 6 we know that at all  $t$  particle conservation requires

$$|\alpha(k; t)|^2 + |\beta(k; t)|^2 = 1, \quad (7.50)$$

leading to the orthonormality (7.49d), see Problem 7.14.

We can now solve the scattering problem. We are seeking the solution  $\phi(x, \tau; t)$  to (7.44) subject to the initial condition (in  $\tau$ ),

$$\phi(x, \xi; t) = \delta(x + \xi), \quad |\xi| > L, \xi < 0, \quad (7.51)$$

for some early value of  $\tau = \xi$ , at all times  $t$ . A long time ( $\tau$ ) in the past, a particle coming from  $x = \infty$  heads towards the potential in the vicinity of  $x = 0$ .

Combining (7.45) for a trial function with (7.48b, 49c) for the frequencies, we can conveniently expand  $\phi(x, \tau; t)$ , when  $x > 0$ , as

$$\begin{aligned} \phi(x, \tau; t) &= \sum_n [A_n(t)e^{-K_n \tau} + B_n(t)e^{K_n \tau}] u_n(x; t) \\ &+ \int_{-\infty}^{\infty} [A(k; t)e^{-i\omega_k \tau} + B(k; t)e^{i\omega_k \tau}] u(x, k; t) dk \end{aligned} \quad (7.52)$$

where the  $A_n, B_n, A$ , and  $B$  are expansion coefficients. In view of the orthonormality of the discrete states as well as their being orthonormal to the scattering states, we get from (7.51, 52):

$$A_n(t)e^{-K_n \xi} + B_n(t)e^{K_n \xi} = u_n(-\xi; t) . \quad (7.53a)$$

Similarly,

$$A(k, t)e^{-i\omega_k \xi} + B(k, t)e^{i\omega_k \xi} = u^*(-\xi, k; t) \quad \omega_k > 0 . \quad (7.53b)$$

The scattering equation (7.44) is, like the classical wave equation, of second order in  $\tau$ . Accordingly, one must fit not only the initial value  $\phi(x, \xi; t)$  but the initial derivative  $\partial\phi(x, \tau; t)/\partial\tau$  to the given conditions at  $\tau = \xi$ , thereby obtaining equations to determine both the  $A, A_n$  and the  $B, B_n$  expansion coefficients, see Problem 7.12. The result is that  $\phi$  for  $x \rightarrow \infty$ , i.e.,  $\phi_\infty$ , is given by

$$\phi_\infty(x, \tau; t) = \sum_n \gamma_n^2(t) e^{-K_n(x-\tau)} + \frac{1}{2\pi} \int_{-\infty}^{\infty} [e^{-ikx} + \beta(k, t)e^{ikx}] e^{-ik\tau} dk . \quad (7.54)$$

(The reduction in Problem 7.13 may be found helpful.)

Let us now recall (7.47):

$$\phi_\infty(x, \tau; t) = \delta(x + \tau) + B(x - \tau; t) \quad (7.47)$$

where, again,  $B$  is the returning signal. If we expand  $\delta(x + \tau)$  in the usual way as

$$\delta(x + \tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ik(x+\tau)} dk , \quad (1.25)$$

we find from (7.54) that

$$B(x - \tau; t) = \sum_n \gamma_n^2(t) e^{-K_n(x-\tau)} + \frac{1}{2\pi} \int_{-\infty}^{\infty} \beta(k, t) e^{ik(x-\tau)} dk . \quad (7.55)$$

The next step is to find the time  $t$  dependence of  $\gamma_n$  and  $\beta$ . In Problem 7.15 we use the wave equation (7.46) and the KdV equation to show that for any discrete eigenstate  $u_n(x; t)$

$$R_n \equiv u_{nt} + v_x u_n - 2(v - 2K_n^2) u_{nx} = 0 , \quad (7.56a)$$

the subscripts  $x, t$  denoting partial differentiation with respect to these variables.

In the asymptotic region  $x \rightarrow \infty, v \rightarrow 0$ , (7.56a) yields for  $\gamma_n(t)$ , [see (7.48a)]

$$\gamma_n(t) = \gamma_n(0) e^{4K_n^3 t} . \quad (7.57a)$$

For the continuum states  $u$  in the same asymptotic region, one finds (Problem 7.15) that if  $k$  in (7.49) is to be independent of  $t$  then

$$R = c(t)u(x, \omega; t). \quad (7.56b)$$

Here  $R$  is  $R_n$  of (7.56a) with  $u$  replacing  $u_n$  and  $-k^2$  replacing  $K_n^2$ . In (7.56b)  $c$  may be any function of  $t$ , but is not a function of  $x$ . Now if the leading term in (7.49a) for  $u$  as  $x \rightarrow \infty$ , namely  $e^{-ikx}$ , is to be an asymptotic solution of (7.56b) we must have

$$c = 4ik^3, \quad (7.58)$$

which also takes care of  $u$  at  $x \rightarrow -\infty$ , (7.49b). With this value for  $c$  we find that if the second term,  $\beta(k; t)e^{ikx}$ , is to be a solution of (7.56b) we must have

$$\beta(k; t) = e^{8ik^3 t}. \quad (7.57b)$$

Finally, (7.57a,b) substituted into (7.55) yields

$$B(x - \tau; t) = \sum_n \gamma_n^2 e^{-K_n(x - \tau) + 8K_n^3 t} + \frac{1}{2\pi} \int_{-\infty}^{\infty} \beta(k) e^{ik(x - \tau) + 8ik^3 t} dk, \quad (7.59)$$

where  $\gamma_n$  and  $\beta(k)$  have their  $t = 0$  values. These values are found by direct calculation of the bound states and scattering states associated with the given  $v(x, 0)$ ; see Example 7.1 and Problem 7.20 for illustration. We notice that  $B(x - \tau; t)$ , which from (7.47) and (7.51) goes to zero as  $\tau \rightarrow -\infty$ , is, in fact, comprised entirely of waves moving to the right in time  $\tau$ , as expected for the scattered part of  $\phi$ .

To reconstruct  $v(x, t)$  from  $\phi(x, \tau; t)$  one starts by representing  $\phi$  by a *linear Fredholm integral equation*:

$$\phi(x, \tau; t) = \phi_{\infty}(x, \tau; t) + \int_x^{\infty} K(x, \xi; t) \phi_{\infty}(\xi, \tau; t) d\xi \quad (7.60)$$

where  $K(x, \xi; t)$  is the kernel of the equation. Substituting this form for  $\phi$  into the original wave equation (7.44), we find in Problem 7.16 that the  $K$  which solves (7.60) will also satisfy

$$K_{\xi\xi}(x, \xi; t) - K_{xx}(x, \xi; t) + v(x, t)K(x, \xi; t) = 0 \quad \xi > x, \quad (7.61a)$$

$$-2 \frac{\partial}{\partial x} K(x, x; t) = v(x; t), \quad (7.61b)$$

$$K, K_{\xi} \rightarrow 0 \quad \text{as} \quad \xi \rightarrow \infty, \quad (7.61c)$$

where the subscripts again denote derivatives. The conditions (7.61) yield a unique solution for the function  $K(x, \xi; t)$ .

Further, going back to the scattering problem, it is clear that no disturbance starting at  $x = -\tau$  and moving in the negative  $x$  direction with unit

velocity will enter the region  $(x + \tau) < 0$  even after scattering. Since  $\phi$  is zero in this region, we can write (7.60) as (causality condition)

$$\phi_\infty(x, \tau; t) + \int_x^\infty K(x, \xi; t) \phi_\infty(\xi, \tau; t) d\xi = 0 \quad x + \tau < 0 .$$

Introducing the expression (7.47) defining the reflected signal  $B$  in terms of  $\phi_\infty$  we have

$$B(x - \tau; t) + K(x, -\tau; t) + \int_x^\infty K(x, \xi; t) B(\xi - \tau) d\xi = 0 \quad x + \tau < 0 .$$

Replacing  $\tau$  by  $-y$ , as is conventional, and  $\xi$  by  $z$ , we obtain

$$K(x, y; t) + B(x + y; t) + \int_x^\infty K(x, z; t) B(z + y; t) dz = 0 \quad y > x . \quad (7.62)$$

Equation (7.62) is referred to as the Gelfand-Levitan-Marchenko equation after its originators. As we shall see, the point is that given  $B$  one can readily solve (7.62), whereas, as mentioned, no direct method is known for solving the KdV equation. Once the integral equation (7.62) is solved, (7.61b) yields  $v(x, t)$ .

We draw attention to the fact that  $B$  in (7.59) consists of two contributions, one originating in the bound states, the other in the continuum states. In Problem 7.17 we give some indication of how the effects of the two contributions can be separated when solving (7.62) for  $K$ . In any case, and recalling the dispersion written into the KdV equation this is not surprising, the effects of the second term quickly die out in time, i.e., as  $t^{-1/3}$  (Problems 7.18 and 7.20), leaving only the effects of the first term, the solitons.

Looking back at (7.59) for  $B$  and ignoring the continuum contribution, we may write (remembering that  $y = -\tau$ )

$$B(x + y; t) = \sum_n g_n(x) h_n(y) \quad (7.63a)$$

$$\text{with } g_n(x; t) \equiv \gamma_n^2 e^{-K_n x + 8K_n^3 t} \quad (7.63b)$$

$$h_n(y) \equiv e^{-K_n y} . \quad (7.63c)$$

Similarly, we choose a trial solution for  $K(x, y; t)$  of the form

$$K(x, y; t) = \sum_n w_n(x; t) h_n(y) , \quad (7.63d)$$

where  $w_n(x; t)$  are functions to be determined. Substituting (7.63) into (7.62) we find

$$\sum_n h_n(y) \left[ w_n(x; t) + g_n(x; t) + \sum_m w_m(x; t) \int_x^\infty h_m(z) g_n(z; t) dz \right] = 0 .$$

Thus, we must have for each  $n$

$$w_n(x; t) + g_n(x; t) + \sum_m w_m(x; t) \int_x^\infty h_m(z) g_n(z; t) dz = 0. \quad (7.64)$$

If there are  $N$  bound states, (7.64) represents a set of  $N$  linear inhomogeneous equations for  $w_m, m = 1, 2, \dots, N$ . In matrix form this set of equations may be written

$$\mathbf{P}\mathbf{W} = -\mathbf{G},$$

$\mathbf{W}, \mathbf{G}$  are column matrices with elements  $w_n, g_n$  respectively, and

$$P_{mn}(x; t) = \delta_{mn} + \int_x^\infty g_m(z; t) h_n(z) dz. \quad (7.65)$$

The solution is

$$\mathbf{W} = -\mathbf{P}^{-1}\mathbf{G},$$

and

$$\begin{aligned} K(x, x; t) &= \mathbf{H}^T(x)\mathbf{W}(x; t) = -\mathbf{H}^T(x)\mathbf{P}^{-1}(x; t)\mathbf{G}(x; t) \\ &= -\sum_{m,n} h_m(x) P_{mn}^{-1} g_n(x; t), \end{aligned}$$

with  $\mathbf{H}^T(x)$  the row matrix having elements  $h_n(x)$ . Since

$$\partial/\partial x P_{mn}(x; t) = -g_m(x; t) h_n(x; t),$$

we also have

$$\begin{aligned} K &= \text{Tr} \{ \mathbf{P}^{-1} \partial \mathbf{P} / \partial x \} \quad (\text{the sum of diagonal elements}) \\ &= \sum_n \sum_m \frac{P'_{mn}}{|P|} \frac{\partial P_{mn}}{\partial x} \\ &= |\mathbf{P}|^{-1} \partial |\mathbf{P}| / \partial x, \end{aligned} \quad (7.66)$$

where  $|\mathbf{P}|$  denotes the determinant of  $\mathbf{P}$ , and  $P'_{mn}$  is the cofactor of  $P_{mn}$ .

We can find  $P_{mn}(x; t)$  in (7.65) from (7.63b,c), which define  $g_n$  and  $h_n$ .

$$P_{mn}(x; t) = \delta_{mn} + \frac{\gamma_m^2 e^{-(K_m + K_n)x + 8K_m^3 t}}{K_m + K_n}. \quad (7.67)$$

Lastly, from (7.61b) relating  $v$  to  $K$ , we obtain

$$v(x, t) = -2 \frac{\partial^2}{\partial x^2} \log |\mathbf{P}(x; t)|. \quad (7.68)$$

To illustrate the use of (7.67) and (7.68) we calculate in Example 7.1 the time evolution of a soliton originating with a delta function wave disturbance. The delta function, it will be remembered from Chap. 6, has only one bound state, leading to a single soliton. It is interesting to see how the initial condition  $v(x, 0)$  determines the soliton evolution. The amplitude of the soliton, as we saw in Sect. 7.4, is directly related to its velocity, an indeed nonlinear result; the soliton comes as a complete package without adjustable amplitude. The number and nature of the solitons are, remarkably enough, respectively determined by the number of bound eigenstates in the initial condition  $v(x, 0)$ , and by the parameters  $\gamma_n, K_n$  characterizing the asymptotic expressions of these states. Any slack in  $v(x, 0)$ , i.e., failure of the solitons to represent  $v$  at the initial time, is taken up via a contribution traceable to the presence of the continuum states in  $B$ ; this additional disturbance then dies out. In Problem 7.19 we find  $v(x, t)$  for two solitons by using (7.67, 68). We thereby capture the rather extraordinary exact solution (7.43) of the previous section. Indeed, (7.67, 68) easily give the exact stable part of the solution to the KdV equation no matter how many solitons may be present.

**Example 7.1.** Find the soliton solution  $v_D(x, t)$  (the D stands for “discrete”) given that at  $t = 0$ ,  $v(x, 0) = -W\delta(x)$ .

*Solution.* In Example 6.2 we learn that the only bound solution to the eigenvalue problem

$$\frac{d^2u}{dx^2} + [\lambda + W\delta(x)]u(x) = 0$$

with  $u$  chosen real, is given by the state

$$\begin{aligned} u(x) &= \sqrt{W/2}e^{-Kx} & x > 0 \\ &= \sqrt{W/2}e^{Kx} & x < 0, \\ \lambda &= -K^2, \quad \text{and} \quad K = W/2. \end{aligned}$$

The eigenvalue equation (7.46) shows that we must identify  $\lambda$  in Chap. 6 with  $\omega^2$  here. Recalling (7.48b), we have

$$\omega \equiv -iK_n \equiv -iK,$$

giving the stated solution for  $u$ .

From (7.48a) for  $\gamma$  and (7.67) for  $P_{mn}$  we find with  $\gamma^2 = W/2 = K$ ,

$$\begin{aligned} |P| &= P_{11} = P = 1 + K \frac{e^{-2Kx+8K^3t}}{2K} \\ &= 1 + e^{-2K(x-x_1)+8K^3t} \end{aligned}$$

with  $x_1$ , the phase shift, given by

$$e^{2Kx_1} = 1/2 \quad \text{or} \quad x_1 = 1/(2K) \ln(1/2).$$

If we let

$$P(x; t) = 1 + e^{-a(x,t)}, \quad a = 2K(x - x_1) - 8K^3t$$

then,

$$\begin{aligned} \partial \ln P / \partial x &= P^{-1} \partial P / \partial x \\ \frac{\partial^2 \ln P}{\partial x^2} &= \frac{-1}{P^2} \left[ \left( \frac{\partial P}{\partial x} \right)^2 - P \frac{\partial^2 P}{\partial x^2} \right] = \frac{-4K^2[e^{-2a} - (1 + e^{-a})e^{-a}]}{(1 + e^{-a})^2} \\ &= 4K^2 \left[ e^{a/2} + e^{-a/2} \right]^{-2}. \end{aligned}$$

Finally, from (7.68) relating  $v$  to  $P$

$$v_0(x, t) = -2K^2 \operatorname{sech}^2[K(x - x_1) - 4K^3t].$$

There will also be a continuum part,  $v_C(x, t)$ , see Problem 7.19.

## Problems

**7.1** Use the method of separation of variables to show that (7.12a,b) is the desired solution of Laplace's equation (7.11) for water waves, subject to boundary conditions (7.4) and (7.10).

*Hints:* Let  $\phi(x, z, t) = U(x)V(z)W(t)$ , with  $W(t) = e^{-i\omega t}$ . Show that the physical solutions of (7.11) have the general form  $U(x) = Ae^{ikx}$  for waves traveling in the positive  $x$  direction,  $V(z) = Ce^{kz} + De^{-kz}$ ,  $k$  real. How must  $C$  and  $D$  be related to satisfy (7.4)? Lastly obtain (7.12b).

**7.2** Use (7.17) to find group velocities for a wave packet describing the propagation of a free particle according to wave mechanics. Give a physical interpretation of the results based, for instance, on the de Broglie relation. Referring to Sect. 6.3, find a suitable choice for " $k_0$ " of (7.17) explaining the role played by the resulting  $s_g$  in describing the propagation, even in the case of dispersion.

**7.3** Show that in shallow water gravity waves are not dispersive at all.

**7.4** Verify (7.20) for  $s_{ph}$  by expanding (7.12b).

**7.5** Verify (7.38) by inserting a trial solution  $v_x = v_{x0}e^{i(kx - \omega t)}$ .

**7.6** Bernoulli's equation for a nonrotating, incompressible fluid is

$$\frac{p(l, t)}{\varrho} + gy + \frac{1}{2}u^2(l, t) = \frac{\partial \phi(l, t)}{\partial t} + c(t),$$

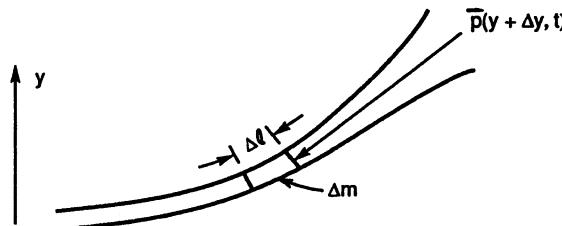
where  $\phi(y, t)$  is the velocity potential,  $p(l, t)$ , is the total pressure, and  $c(t)$  is a function of  $t$  only. The force on the fluid element  $\Delta m$  along the tube will be

$$A [p(l + \Delta l, t) - p(l, t)] ,$$

along with the influence of gravity.

a) Use Newton's second law, taking full cognizance of nonlinearity, to obtain the equation.

b) What form does the equation take for steady state?



**Fig. 7.6.** Problem 7.6.  
A non-viscous fluid confined to a frictionless tube. The vertical direction is given by the  $y$ -axis

**7.7** a) Show that the wave equation (3.3) for waves on a stretched string can be factored to give the two independent first order pdes

$$\left( \frac{\partial}{\partial t} + s \frac{\partial}{\partial x} \right) y_1(x, t) = 0 ,$$

$$\left( \frac{\partial}{\partial t} - s \frac{\partial}{\partial x} \right) y_2(x, t) = 0 .$$

b) Find the characteristics for these two first order pdes.

c) i) If the initial condition for the total wave  $y(x, t)$  satisfying (3.3),  $y(x, 0)$  is given, are  $y_1(x, t), y_2(x, t)$  completely determined?

ii) How do we incorporate a second initial condition, namely,  $dy(x, t)/dt|_{t=0} = 0$  for all  $x, L > x > 0$ ?

iii) If we know  $u(x, 0)$  for all  $x, x_F > x > x_B$  in Sect. 7.3, is this sufficient to find  $u(x, t)$  at all subsequent times? Discuss.

d) Use the characteristics found in (b) to divide the  $xt$  plane into six regions as is done in Sect. 7.3. Explain what happens in each region given that

$$\begin{aligned} y(x, 0) &= \text{given function} & x_F \geq x \geq x_B \\ &= 0 & x > x_F, x < x_B \\ \frac{\partial y(x, t)}{\partial t} \Big|_{t=0} &= 0 & \text{for all } x . \end{aligned}$$

**7.8** a) Linearize (7.23, 24), and show that this yields  $s_0^2 = gd$ , i.e., (7.19).

b) Recalling (7.6) giving the excess pressure as a function of the surface displacement  $h(x, t)$ , use the equations in (a) to show

- i) that  $h = du/s_0$ ,
- ii) that for shallow waves,  $\partial h / \partial t$  is indeed small compared to  $u$ , again justifying the one dimensional treatment in Sect. 7.3. However, if we need  $h$ , we can find it approximately from (i).

**7.9** Show that in Region IV of Fig. 7.3 we have both  $u(x, t) = P(x, t)$  and  $u(x, t) = -P(x, t)$ . What conclusion do you draw?

**7.10** Show that, indeed,  $v(x, t)$  as given in (7.41):

$$v(x, t) = -\alpha^2/2 \operatorname{sech}^2 [\alpha/2(x - \alpha^2 t - x_1)]$$

is a solution of the KdV equation (7.40).

**7.11** a) Show that the expression (7.43) for two solitons reduces to (7.41) describing one soliton when  $f_2$  is set equal to zero.

b) Show that in regions of  $x, t$  when  $f_1 \simeq 1$  and  $f_2$  is either *large* or small, that  $v(x, t)$  is given almost exactly by the single  $\alpha_1$  soliton. Is this a reciprocal relationship as regards  $f_1$  and  $f_2$ ? If so, how do we know?

c) Show that for  $t \rightarrow -\infty$   $f_2 \ll 1$  when  $f_1 \simeq 1$ , and that  $f_1 \gg 1$  when  $f_2 \simeq 1$ . Assume  $\alpha_2^2 \gg \alpha_1^2$ . Which wave is leading at  $t \rightarrow -\infty$ ?

d) Now show that as  $t \rightarrow \infty$  when  $f_1 \simeq 1, f_2 \gg 1$ , and that when  $f_2 \simeq 1, f_1 \ll 1$ .

e) How do the above facts support the picture of two solitons passing each other without change of shape? Show that for soliton 2 the effect of the collision is a forward shift in space of  $(2/\alpha_2) \log[(\alpha_2 + \alpha_1)/(\alpha_2 - \alpha_1)]$ , and for Soliton 1 a backward shift of  $2/\alpha_1 \log[(\alpha_2 + \alpha_1)/(\alpha_2 - \alpha_2)]$ .

**7.12** Show that by fitting  $\partial\phi(x, \tau; t)/\partial\tau|_{\tau=\xi}$  to  $\partial\delta(x+\tau)/\partial\tau|_{\tau=\xi}$  along with (7.53b) determines  $A_n(t), B_n(t), A(k; t), B(k; t)$ . Find these coefficients, the latter two for both positive and negative values of  $k$ .

**7.13** Show that at  $x \rightarrow \infty$

$$\begin{aligned} & \int_0^\infty [e^{-ikx} + \beta(k)e^{ikx}] [e^{-i\tau k} + \beta^*(k)e^{i\tau k}] dk + |\alpha(k)|^2 \int_{-\infty}^0 e^{-ikx} e^{-i\tau k} dk \\ &= \int_{-\infty}^\infty [e^{-ikx} + \beta(k)e^{ikx}] e^{-ik\tau} dk \end{aligned}$$

by virtue of (7.50), and the fact that  $\beta^*(k) = \beta(-k)$ . This last equality follows from the invariance of the theory to whether we use  $i$  or  $-i$ , the scattering equation (7.44) being linear in  $\phi$  and with real coefficients.

Use the result of this problem along with those of Problem 7.12 to obtain the expansion for  $\phi_\infty$  in (7.54) from (7.52).

**7.14** Check that as  $n \rightarrow \infty$

$$\begin{aligned} & \frac{1}{2\pi} \int_0^\infty e^{-x^2/4n} [e^{-ikx} + \beta(k)e^{ikx}] [e^{ikx} + \beta^*(-k)e^{-ikx}] dx \\ &+ \frac{1}{2\pi} \int_{-\infty}^0 e^{-x^2/4n} [\alpha(-k)e^{-ikx}] [\alpha^*(-k)e^{ikx}] dx = g_n(0), \end{aligned}$$

and explain the significance of this calculation. What happens to the contribution to the integral from the region  $L > x > -L$ , where  $v(x, t) \neq 0$ , if we consider the actual eigenstates  $u(x, k; t)$  rather than their asymptotic forms? What happens if  $n/L \rightarrow \infty$ ?

**7.15** a) By differentiating the eigenvalue equation (7.46) with respect to both  $x$  and  $t$ , and invoking the KdV equation for  $v(x, t)$ , show that

$$\frac{\partial}{\partial x}(u_x R - u R_x) = \frac{\partial \omega^2}{\partial t} u^2 \quad (7.69)$$

where  $u_x = \partial u(x, \omega(t); t)/\partial x$ , with  $R_x$  similarly a derivative of  $R$ :

$$R \equiv u_t + v_x u - 2(v + 2\omega^2)u_x .$$

b) Integrating both sides of (7.69) with respect to  $x$  from  $-\infty$  to  $\infty$ , show that  $\partial \omega^2 / \partial t$  must vanish for bound states that go to zero as  $x \rightarrow \pm\infty$ , the eigenvalues are constant in  $t$ .

c) Using the results of (a) and (b), show that when  $u = u_n$ , giving

$$R \equiv R_n , \quad R_n = 0 . \quad (7.56a)$$

*Hint:* Integrate twice over  $x$  from  $-\infty$  to  $x$ .

d) Show that for free states  $u(x, \omega; t)$  a consequence of (7.69) is that if the eigenvalues  $k^2(t)$  are to be independent of  $t$  then

$$R = c(t)u(x, \omega; t) , \quad (7.56b)$$

where  $c$  depends on  $t$  but not  $x$ .

**7.16** Show that  $\phi(x, \tau; t)$ , or temporarily dropping the  $t$ ,  $\phi(x, \tau)$  is indeed a solution of the wave equation (7.44),

$$\phi_{xx} - \phi_{\tau\tau} - v\phi = 0 \quad (7.44)$$

if  $K$  in

$$\phi(x, \tau) = \phi_\infty(x, \tau) + \int_x^\infty K(x, \xi)\phi_\infty(\xi, \tau)d\xi \quad (7.60)$$

satisfies

$$K_{\xi\xi} - K_{xx} + v(x)K = 0 \quad \xi > x \quad (7.61a)$$

$$v(x) = -2\frac{\partial}{\partial x}K(x, x) \quad (7.61b)$$

$$K, K_\xi \rightarrow 0 \quad \xi \rightarrow \infty .$$

The subscripts denote derivatives.

*Hint:* What is the differential equation satisfied by  $\phi_\infty$  in view of the fact that  $v(x) \rightarrow \infty$  as  $x \rightarrow \infty$ ? When substituting  $\phi(x, \tau)$  into the wave equation use integration by parts repeatedly to evaluate  $\phi_{\tau\tau}$ .

**7.17** Show that the integral equation (7.62) can be written as

$$(1 + \mathbf{b}_D)K + \mathbf{b}_C K = -(B_D + B_C),$$

where we have split  $B(x - \tau)$  in (7.59) into its “discrete part”  $B_D$  and its “continuous part”  $B_C$ , and the operator  $\mathbf{b}_D$  is defined by

$$\mathbf{b}_D K \equiv \int_x^\infty K(x, z) B_D(z + y) dz.$$

Show also that if we split  $K$  into

$$K \equiv K_C + K_D,$$

then

$$(1 + \mathbf{b}_D)K_D = -B_D,$$

$$(1 + \mathbf{b}_C)K_C = -B_C,$$

providing the cross terms  $\mathbf{b}_D K_C$  and  $\mathbf{b}_C K_D$  may be left out. This approximation becomes increasingly valid as  $t \rightarrow \infty$ , then allowing for the separation of  $v(x, t)$  into solitons coming from  $K_D$ , and a time decaying wave disturbance originating with the continuum.

**7.18** Solitons aside (where the procedure would not converge) the KdV equation can be solved directly by iteration.

a) Show that the KdV equation (7.40) can be written as

$$Lv = 3\partial/\partial x(v^2),$$

where

$$Lv = \partial v / \partial t + \partial^3 v / \partial x^3.$$

b) Show that the expansion

$$v = v_1 + v_2 + \dots$$

leads to the iterative scheme

$$Lv_1 = 0$$

$$Lv_2 = -3\partial/\partial x v_1^2$$

...

$$Lv_n = -3 \sum_{j=1}^{n-1} \partial/\partial_x l(v_j v_{n-j}),$$

and has the solution

$$f(x, t) = \int_{-\infty}^{\infty} \phi(x') h(x - x', t) dx' + \int_{-\infty}^{\infty} \int_0^t \varrho(x', t') h(x - x', t - t') dx' dt'.$$

*Hint:* Refer to Problem (1.17) to find  $h(x - x', t)$ .

c) A well known integral yields

$$h(x, t) = (3t)^{-1/3} A_i \left[ x(3t)^{-1/2} \right].$$

Assuming that the Airy function  $A_i$  is bound and decays quickly for a finite argument, make plausible, on the basis of the foregoing two problems (7.17, 18a,b,c), that as stated in the text the nonsoliton part of  $v$  decays as  $t^{-1/3}$  for any  $x$ . More complete insight can be obtained by looking up H. Segur, J. Fluid Mech. **59**, part 4, 721 (1973).

**7.19** a) Use (7.67) to show that for two solitons

$$|P| = 1 + f_1 + f_2 + \Lambda f_1 f_2,$$

where  $f_1$  and  $f_2$  are defined via (7.42a), with  $\alpha_1 = 2K_1$ ,  $\alpha_2 = 2K_2$ ,  $\gamma_1^2 = \alpha_1 e^{\alpha_1 x_1}$ ,  $\gamma_2^2 = \alpha_2 e^{\alpha_2 x_2}$ , and  $\Lambda = (\alpha_1 - \alpha_2)^2 / (\alpha_1 + \alpha_2)^2$ .

b) Use (7.68) giving  $v(x; t)$  in terms of the determinant  $|P(x; t)|$  to generate (7.43), the disturbance produced by two interacting solitons. Is the result an exact solution of (7.40), the KdV equation?

**7.20** a) Find the scattering states for the initial disturbance considered in Example 7.1,

$$v(x, 0) = -W \delta(x).$$

b) Calculate the contribution of the these states to  $B(x, t)$ .

## Further Reading

T. Erber, K.A. McGreer, E.R. Nowak, J.D. Wan, H. Weinstock: “Onset of hysteresis measured by scanning tunneling microscopy”, J. Appl. Phys. **68**, 1370 (1990)

J. Lighthill: *Waves in Fluids* (Cambridge University Press, 1978)

The full treatment of waves on water, linear and nonlinear, can be found in this treatise.

J.A. Krumhansl: “*Unity in the Science of Physics*”, Physics Today, **33**, March 1991.

The reader will find here a very helpful overview of the increasing importance of solitons in physics.

G.B. Whitham: *Linear and Nonlinear Waves* (Wiley, New York 1974)

In the sections on solitons the method of inverse scattering due to G.N. Balanis is described in Whitham’s comprehensive treatise. (G.B. Whitham was a professor of G.N. Balanis.)

P.G. Drazin, R.S. Johnson: *Solitons: an Introduction* (Cambridge University Press, 1989)

A good source to look at for the reader who wishes to know more about the many aspects of the theory of solitons on an intermediate level.

M. Remoissenet: *Waves Called Solitons, Concepts and Experiments* (Springer-Verlag, Berlin Heidelberg 1996), 2nd edition.

## 8. Nonlinear Phenomena – Chaos

**Preamble.** Chapters 1–6 have dealt with linear behavior, behavior that forms the core of traditional theoretical physics. The last two chapters, Chaps. 7, 8, give the reader a glimpse into an extension of the traditional, – the currently active field of nonlinear behavior. In this chapter two essays by distinguished physicists deal with order and chaos, essays complementary to the subject matter of Chap. 7, which revealed unexpected order and stabilities.

The authors of the first contribution, Andrei V. Gaponov-Grekhov and Mikhail I. Rabinovich have been working for decades on problems in nonlinear sciences at the Institute of Applied Physics of the Russian Academy of Sciences in Nizhnii Novgorod (Gorkii). Gorkii, known as the exile of the late A.D. Sakhorov, was until shortly closed to foreigners; scientific work done at its institute has taken unusual time to reach the international community of scientists. Professor Gaponov-Grekhov is a member of the Russian Academy of Sciences, and Professor Rabinovich a corresponding member.

The essay of Gaponov-Grekhov and Rabinovich is an abridged version of Chap. 6 of their textbook *Nonlinearity in Action* released by Springer-Verlag in 1992. Their chapter has been shortened for this text by Dr. Ernst F. Hefter (Springer-Verlag). In the original, colored illustrations exhibit the unexpected beauty that comes with the interplay between chaos and order.

The second essay has been prepared by Martin C. Gutzwiller. It is essentially a reprint of his review *Quantum Chaos* that appeared also with beautiful colored illustrations in *Scientific American* in January 1992. Gutzwiller, a member of the US National Academy of Sciences, is a staff member at the IBM research laboratory in Yorktown Heights, New York, and also a professor at Columbia University. He is the author of a graduate level textbook “Chaos in Classical and Quantum Mechanics”, recently published by Springer-Verlag.

## 8.1 Nonlinear Physics – Chaos and Order<sup>1</sup>

By Andrei V. Gaponov-Grekov and Mikhail I. Rabinovich

### The Where and the How

In science as in daily life we run into questions the answers to which are so sophisticated and uncertain that it appears most reasonable to adopt either the point of view established by the majority (or convention) or to treat them not as a matter of science but of belief. Into the latter category fall two questions that are possibly some of the oldest ones: “Where does randomness come from?” and “How does order arise?” Everybody who has pondered about the basic principles lying at the foundations of Nature must have encountered these questions. From daily life as well as from traditional learning we have grown accustomed to the almost obvious perception that complicated, irregular, imbroglio behavior is possible only in very complicated systems. Examples are the vast number of molecules in a balloon filled with gas or a crowd of excited football fans just after the unexpected announcement that *the* football game of the season is cancelled. In the case of such complicated systems we are usually not able to unravel for the individual events unambiguous connections between cause and consequences, that is, we are not in the position to predict the detailed behavior of the system and thus consider it to be ruled by chance.

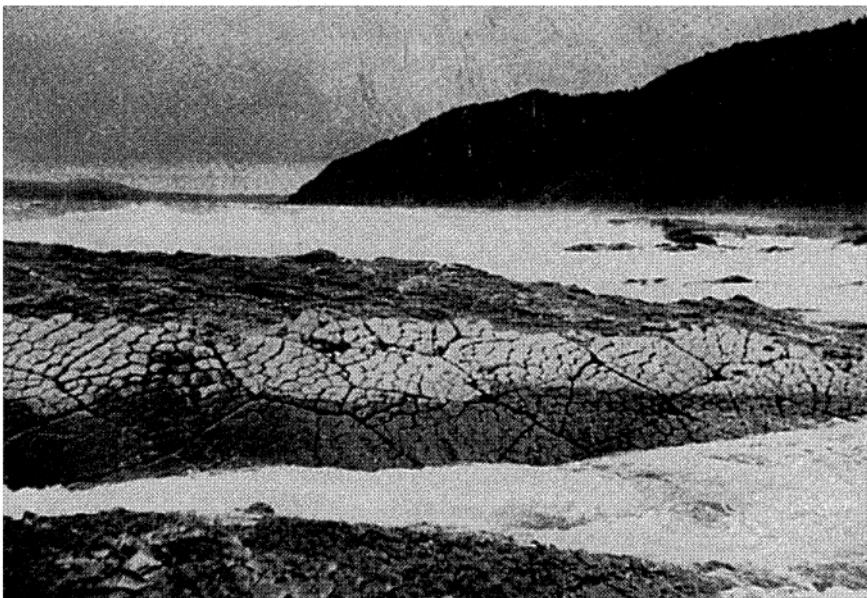
It is true that scientists always nourished the hope that in principle such an interpretation in terms of random behavior and our inability to predict the future of the system could be removed once we had a more complete knowledge of its details. For a long time the following point of view was accepted: if we obtain more precise knowledge on the details of the interaction of all elements of a complicated system and collect more detailed information on their initial conditions then we will be able to predict its behavior over long periods and the randomness “will become smaller and smaller” as we increase our knowledge.

Similarly natural appears the idea that there must be an “organizing beginning” or “creator” behind each sophisticatedly organized structure which exists in a stable form or which emerged out of a background of noise and disorder. It is precisely for these reasons that the strongly regular structures of clouds or the hexagons of volcanic origin depicted in Fig. 8.1 appear so imbroglio and even mystic.

In this short excursion we try to get a glimpse of the beauty of the most intriguing facets of modern nonlinear dynamics, *disorder* versus *ordered structures*. Chaos and order do not just coexist in Nature, they arise due to

---

<sup>1</sup> The more detailed original appeared as Chap. 6 of the book “Nonlinearities in Action” by the same authors (Springer, Berlin, Heidelberg, New York 1992)



**Fig. 8.1.** Hexagonal volcanic structures (Kuril island “Kunashir”) [illustration courtesy of E.N. Pelinovsky; Inst. Appl. Phys., Russian Acad. Sci., Nizhnii Novgorod, Russia]

the same general principles, laws, or conditions which justifies to discuss them in parallel.

In the last few decades two distinct discoveries have been made that have completely changed our perception of the nature of randomness and regularity. It has turned out that very simple systems are able to exhibit random behavior. These are systems evolving or living according to very simple rules (or having a small number of elements or degrees of freedom). Randomness is an intrinsic property of such a system and is not imposed on it by its surroundings or through external forces. It is impossible to remove it through more detailed studies of the system. Such a randomness exhibited even by simple regular systems is what we refer to as *dynamical chaos*.

The other discovery is related to the awareness and experimental confirmation of the fact that from initial disorder simple as well as complicated highly organized structures may spontaneously arise, continuously developing and evolving. This process is referred to as *self-organization*.

### Randomness Born out of Nonrandomness

Let us start with chaos. The main difficulty which we have to overcome is a psychological one. We are used to the notion that simple systems (like a swing, a marble in a chute, etc.) display a very simple behavior. Knowing

the rules such a simple system obeys and its initial conditions, we may predict its behavior over periods of time of any desired length. For example, using Newton's laws we can list all future solar eclipses not only for the next few hundred but even thousand years! We need relatively little time to evaluate them, hence, this is indeed a real prediction.

What are then the sources within similarly simple systems – systems living according to simple rules – which give rise to unpredictability, irregularity, randomness? To develop an understanding of such peculiar features, let us start with a simple example.

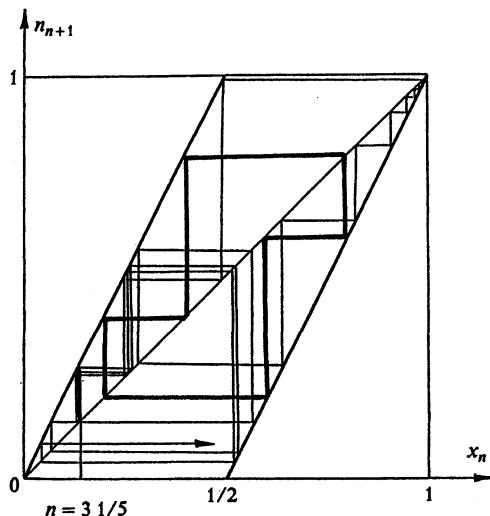
Presumably all of us encountered back in our schooldays sequences of numbers like the arithmetic series  $x_{n+1} = x_n + a$  and the geometrical series  $x_{n+1} = bx_n$ . Let us refer to these series as systems. In this interpretation the label or number of each member or element may be viewed as a time (measured in discrete units) and  $x_n$  as the state of the system at the moment  $n$ . The behavior of such simple systems is indeed very simple, because their states are completely predictable for as long as you like. For arbitrarily large  $n$  we have  $x_{n+1} = x_1 + na$  and  $x_{n+1} = b^n x_1$ , respectively. Thus a knowledge of the systems at  $n \rightarrow \infty$  does not require that we explicitly know its states at all intermediate moments of time in their “life”. It is sufficient to evaluate  $x_{n+1}$ . Is that always true?

Let us consider a series which is almost equally simple and defined by the rule  $x_{n+1} = \{2x_n\}$  where the symbol  $\{\dots\}$  stands for taking only the fractional part of the number within these brackets. Obviously the  $x_n$  represent points in the interval  $(0,1)$  from 0 to 1 of the number axis. As an example we take for  $x_1$  the number  $1/5$ . Whence follow  $x_2 = 2/5$ ,  $x_3 = 4/5$ ,  $x_4 = 3/5$ ,  $x_5 = 1/5 = x_1$ . Indeed, everything is simple: we observe a purely regular motion and for  $x_1 = 1/5$  we are able to predict the state of the series after arbitrarily long periods of “time”, that is for large  $n$ .

Let us give it another try. As the initial number we now take the fractional part of  $\pi$ . With a pocket calculator you can easily verify that with this we get  $x_1 = \{\pi - 3\}$  and a chain  $x_2, x_3, \dots, x_n$  which never closes and that the resulting sequence of numbers is similar to one of random numbers. We are no longer able to predict the state of the system. Not even for a few steps ahead. It is now impossible to skip intermediate steps without losing the ability to predict the “future”. To give the explicit value of  $x_{n+1}$ , we have to evaluate all  $x_n$  up to the  $(n+1)$ -th step. Since the computational time is of the same order as the time of the moment the prediction is made for, that implies unpredictability (chaos) which is characteristic for such a random series. The same will hold for any series if its first element  $x_1$  is an irrational number. And the overwhelming majority of numbers in the interval from 0 to 1 is of this type.

As illustrated in Fig. 8.2, the motion of our simple system may also be represented graphically. The two axes represent  $x_{n+1}$  and  $x_n$ , respectively. The area in between the lines  $x_{n+1} = 2x_n$  and  $x_{n+1} = 1 + 2x_n$  within the

interval  $(0, 1)$  is the graphical representation of the rule: extend the initial bit by two and take away the integer. Having reflected the trajectory off the bisectrix it converts the final result of the preceding step into the initial value of the following. The closed trajectories (bold lines) correspond to periodic motions of the system. The trajectories (thin lines) which never close are examples of complicated, chaotic motions.



**Fig. 8.2.** Periodic (bold lines) and random (thin lines) trajectories in the simple system defined by  $x_{n+1} = \{x_n\}$

For computer calculations it is convenient to represent  $x_n$  in a binary system. In that case each subsequent value in the series is obtained from the preceding one simply by moving the number one digit to the left (which corresponds to multiplication by two) and by suppressing its integer part. For example, if  $x_1 = 0.101100010111\dots$  then we get  $x_2 = 0.01100010111\dots$ ,  $x_3 = 0.1100010111\dots$  etc. Periodic sequences correspond to rational numbers and random series to irrational numbers. (Let us recall that the irrational numbers are in the majority, implying that almost all arithmetic sequences should be random; usually we do not pay attention to this point because we use normally rational numbers.)

From these examples we have seen that simple fully deterministic systems may indeed give birth to chaos.

### An Unstable Path and Steady Motion – Are They Incompatibles?

The points  $x_n, x_{n+1}, \dots$  wander in a random manner within the interval  $(0,1)$  extending from 0 to 1. Yet, this chaos does have its own order. To see

that in a more pictorial way, let us go over from a series of real numbers to one with complex numbers. As an example we take the series represented by  $x_{n+1} = x_n^2 + C$ . The majority of the resulting points will no longer be found within the interval  $(0,1)$  but in the plane defined by the real and imaginary parts of  $x_n$  namely  $\operatorname{Re} x_n$  and  $\operatorname{Im} x_n$ , respectively. For  $n \sim 10^6$  and  $C = 0.32 + i0.43$ ;  $\operatorname{Re} x = (-2.0 \dots + 2.0)$ ;  $\operatorname{Im} x = (-1.8 \dots + 1.8)$  this leads to pictures of unusual beauty (see the books by Peitgen and Richter and by the authors of this essay).

Each one of the points of this series with the coordinates  $\operatorname{Re} x_1, \operatorname{Re} x_2, \dots, \operatorname{Re} x_n, \dots$  and  $\operatorname{Im} x_1, \operatorname{Im} x_2, \dots, \operatorname{Im} x_n, \dots$  is random. Yet, it is problematic to refer to this manifold as being random. For example, it resembles only vaguely the random figure formed when a handful of sandgrains is thrown at random to the ground. The surprising beauty of this figure is, as a matter of fact, a consequence of the point that the random series emerged in the given case from a deterministic system. It has its own order, that is the dynamical chaos is in a specific way “organized”. This is a remarkable phenomenon which demands a more detailed discussion. However, first we have to answer the question left out earlier: amongst the manifold of series’ the majority of which are random, there are also periodic ones; maybe they are the only ones that are indeed observed?

No. They are unstable, implying that for arbitrarily small uncertainties or variations in the given  $x_1$  we will already have a completely different series, see Fig. 8.2. Here we note a manifestation of another feature of dynamical chaos, an extremely strong sensitivity to changes in the initial conditions. Thus, in spite of the fact that there are many periodic series, even infinitely many (yet, there are nevertheless many, many more nonperiodic ones) it is practically impossible to observe them. However, the nonperiodic trajectories are also unstable. Why is it then that they are observed at all? The point is that there is an almost continuous manifold of them and that they are indistinguishable. Indeed, let us assume that because of a small perturbation (or uncertainty in the given  $x_1$ ) we obtained not the “correct” series, nevertheless the other one which we obtained belongs also to the same manifold of unstable trajectories. Thus we will always observe one or the other of them. Consequently we are led to a remarkable conclusion: although the individual trajectories of a stochastic manifold are unstable and thus not observable (not realized), the manifold itself is stable and at least one of its many trajectories (covering almost the entire manifold) is observed!

### Does Chance Rule the World?

Why didn’t we realize earlier the random behavior of nonrandom systems? Maybe, we were only picking artificial examples? Maybe life is different and real systems with chaos are described by different equations? For example, could it be that Newton’s equations guarantee only a regular behavior of mechanical systems?

Let us consider a very simple system, a swing. Of course, it is described by the simple equations of mechanics which some of us encountered at school. But apparently the example is not a good one: as we know, the swing oscillates periodically; however, its period is also determined by the actions of the person using it, standing up or squatting. The effective length of the swing may thus be altered to accelerate it or (by getting up out of phase) to slow it down. Let us now deprive the person on the swing of his will power. Let us assume that it is now a robot who squats and stands up strictly periodically! What will happen? – The swing will oscillate completely randomwise, without any random forces acting on it! Squatting strictly periodically the robot will out of phase “lengthen” or “shorten” the swing and consequently add energy to the swing (or pendulum) or take it away, respectively. The observed randomness is due to the fact that the swing moves with different speeds depending on its angle of deviation from the vertical. For example, close to the position “head over heels” it slows down drastically and at the turning points it even comes to a stop. Thus we see that even real mechanical systems that obey Newton’s laws, behave chaotically. That is, they generate randomness! Why didn’t we observe it earlier? It seems appropriate to respond: “We saw it, but we did not realize what we were seeing.” The traditional ways of thinking did not allow us to take serious individual experiments demonstrating chaos in simple systems (and therefore not fitting into the framework of established theories). In some way or the other they were “explained” in terms of natural fluctuations or by the salutary influence of noise that could not be accounted for in the calculations.

The sensitivity to the initial conditions typical for chaos is due to the instability of the motion. In a very pictorial and direct way reflected by historical developments.

History would have a very mystic character, if randomness played no role in its evolution. However, during different periods of time in the evolution of society and in different countries this randomness has manifested itself differently. In periods of *stable* evolution, randomness (like the death or murder of a political leader, natural calamity, etc.) pushed the evolutionary path of the society only from one trajectory to another one lying in its immediate neighbourhood: the welfare of the people changed a bit, building plans were altered, interior or exterior politics came to be harsher or more moderate. A qualitatively different picture is observed in a period of *unstable* evolution (just before the outbreaks of wars or revolutions, in periods of disorder and unrest): small random deviations have led to completely different routes in the subsequent development of the society. In such periods it appears indeed as if “chance rules the world”. But in such a case one usually forgets that such significant results due to the action of randomness are in reality only possible because the “flow of life” which was completely changed by an insignificant randomness was unstable and randomness just played the role of a trigger.

The following children's rhyme illustrates such a situation neatly:

For the want of a nail, the shoe was cast.  
 For the want of a shoe, the horse was lame.  
 For the want of a horse, the message was late.  
 For the want of a warning, the city was lost.

This is a pictorial example for a motion that is at every point unstable: if there were a nail, then the horseshoe would have been fixed; if the horseshoe were fixed then the horse would not have been lame; if the horse were not lame (or if there were a spare one) then the message would have been in time; if the message were received in time then the city would not have been lost.

### What is the Character of Nature? – Integer or Fractal?

Pictures due to fractal manifolds of points depicted in the plane of random sequences generated by a deterministic system are remarkably regular see (Plate 1 of) the book by the authors or the one by Peitgen and Richter which contains a multitude of such manifolds with microstructures. According to the words of Ruelle they possess an “aesthetic attractiveness”. These systems of curves, these clouds of points resemble sometimes fireworks of galaxies, sometimes strange mysterious brushwood. This is a sphere of research in which new harmonies will be discovered. Such manifolds with “floating” rarefying structures (which have as a rule noninteger dimensions as will be explained below) are called *fractals*.

An example of fractals is given by the sophisticated evolution of the discrete analogue or predator-prey system modelled by the Lotka-Volterra equations

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} x_n \\ y_n \end{pmatrix} + \frac{h}{2} \begin{pmatrix} f(x_n, y_n) + \varrho f(x_n, y_n) y_n + \varrho g(x_n, y_n) \\ g(x_n, y_n) + g[x_n + \varrho f(x_n, y_n) + \varrho g(x_n, y_n)] \end{pmatrix}.$$

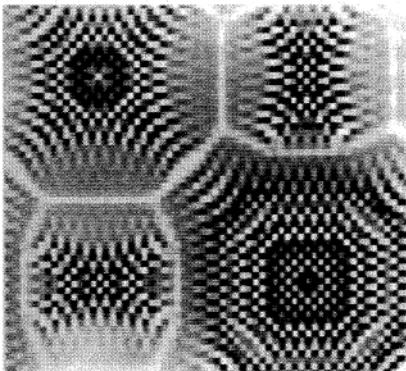
Another one is the stochastic manifold of the nonautonomous self-oscillations of a system like

$$\ddot{x} + \gamma \dot{x} + x + x^3 = F_0 \sin(\omega t).$$

The discrete lattice described by

$$\begin{aligned} \frac{du_{jl}}{dt} &= u_{jl} - (1 + i\beta)|u_{jl}|^2 u_{jl} \\ &\quad + \kappa(1 - ic)(u_{j,l+1} + u_{j,l-1} + u_{j+1,l} + u_{j-1,l} - 4u_{jl}) \end{aligned}$$

with  $j, l = 1, 2, \dots, N$  and  $N \gg 1$  yields for a large supercriticality, that is for approximately  $r = 1/\kappa \geq 0$ , to spatio-temporal chaos. In this process the short-wave excitations play the most important role. In the extreme case the system leads to so-called  $\pi$ - $\pi$ -oscillations corresponding to anti-phase oscillations of neighbouring lattice elements. The arising chaotic structure is very fascinating – it is a kaleidoscope of pictures that never repeat in



**Fig. 8.3.** Modulation waves and  $\pi$ - $\pi$ -oscillations in the discrete lattice-model

time. The complexity of its spatial structure is determined by the number of elements in the lattice, see Fig. 8.3.

Although the term fractal was rather recently introduced into science by Benoit Mandelbrot, fractal objects have been the subject of research for quite a long time. Judging from their huge number, Nature “likes” fractal forms very much. Amongst them we have colloids, electrical discharges, porous solid states, coastal lines, the structures of turbulent flows (the majority of whose trajectories are points generated by deterministic systems with chaotic behavior) and many more. No doubt, the beauty of fractals is linked to their property of self-similarity which manifests itself in repeated patterns at changing scales of the figures, see Peitgen and Richter.

If we magnify under a microscope an arbitrarily small area of a fractal (or stochastic) manifold, i.e. if we increase the resolution significantly, then we observe again an exceptionally complicated picture with a manifold of details. And thus we may proceed to infinity! – a most extraordinary fact. Really up to infinity? What determines the threshold for this chain of subsequent reductions and detailizations? Before answering this question, let us consider what could possibly perturb such a process. Let us recall the underlying rule: what we need is that in between two arbitrary trajectories (for example, strange attractors), we can find as many further trajectories as we like (!) – no matter how close our initial two trajectories are to each other. Obviously this is only possible if the space in which they move is continuous!

It is the discreteness of space (if it corresponds to the given problem) that is the cause for limitations in subsequent detailizations of fractal structures. And for dynamical chaos it is the cause of the existence of chaos itself. Let us look into this in more detail.

Let us imagine space to be discrete (no matter whether it is the real space in which liquid particles move or an abstract state space or phase space). This discreteness may be caused by different reasons. In the case of the flow of a liquid this is naturally related to the impossibility of considering the path

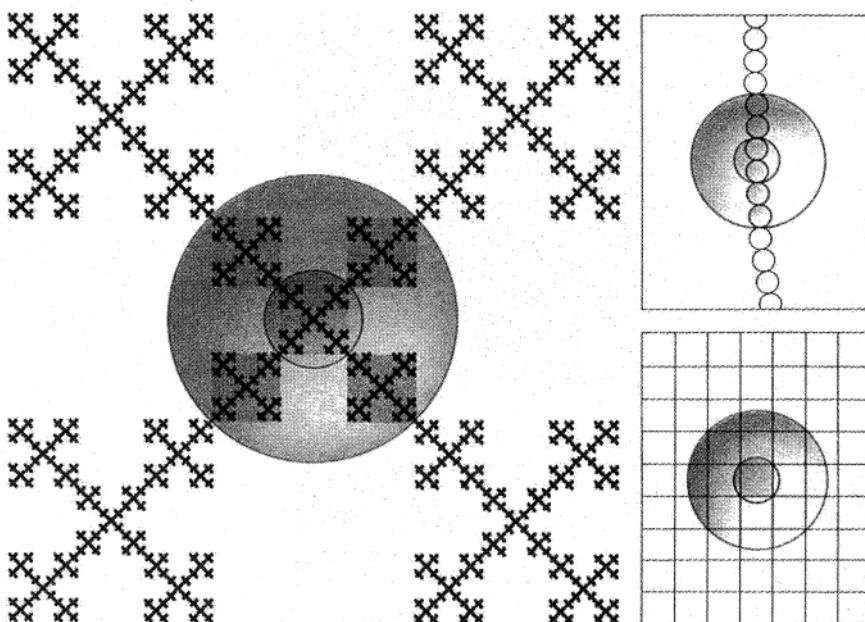
of a fluid particle to be an infinitely small line. In that case the trajectory of each particle will be closed and chaos disappears.

The discreteness of space may also be determined by deeper reasons. In the first place, it may be related to the discreteness of the elements which make up the flow of a liquid or gas (for example, the molecules) in real space. It may also be related to the quantum nature of the objects under investigation because in that case the quantization condition gives rise to a discrete phase space.

### Fractal Fingers

In passing, we mentioned that fractal manifolds have noninteger dimension. What does that mean? Maybe, noninteger dimensions are the result of approximations? It is just the opposite. Estimating the dimensions of an object by eye we make them due to our habits integer ones. Indeed, what is the dimension of an entangled “ball” or clot of thin thread? If we look at it from the far distance we notice a object implying zero dimension. If we come closer then its dimension is three. If we want to consider this object in detail and follow the individual bits of thread, the tiny building units which make up this clot then its dimension seems to be unity. But why does the dimension change in jumps? Well, because this is what our perception is used to and, consequently, it *appears* to us to be true. To obtain the correct answer in the analysis of such unusual objects like fractals (to which, in a crude way, our ball or clot also belongs) we have to define the term dimension more rigidly.

This may be accomplished by counting the number of elements of the object of interest found within a sphere of radius  $r$  with its center within the object under consideration, see Fig. 8.4. The number  $N$  of elements found within such a sphere will be proportional to the  $D$ -th power of the radius of the sphere with  $D$  standing for the dimension of the system. That is, using unity for the associated dimensional constant we have  $N = r^D$  or  $D = \ln N / \ln r$ . For an object in the form of a straight line the number of its elements within such a sphere increases by a factor of three if the radius is tripled, that is  $D = 1$ . If the object is the close packing of elements in a plane then tripling  $r$  implies that  $N$  is increased nine times. Consequently we obtain the dimension  $D = 2$ . So far the rigorous definition gives only the obvious traditionally expected results. However, let us turn to the less trivial situation depicted in Fig. 8.4. For this structure tripling the radius  $r$  of the sphere leads to five times more elements within the sphere, i.e. the number of elements grows faster than for one-dimensional object and slower than for a two-dimensional one. This is related to the fact that our structure is everywhere “porous”, “holey” or fractal so that its dimension turns out to be a noninteger one, namely  $D = 1.46$ .



**Fig. 8.4.** A fractal “snowflake” [courtesy of T. Vicsek of the Research Institute for Technical Physics of the Hungarian Academy of Sciences in Budapest and of L.M. Sander; *Scientific American* **256**, 82 (1987)]

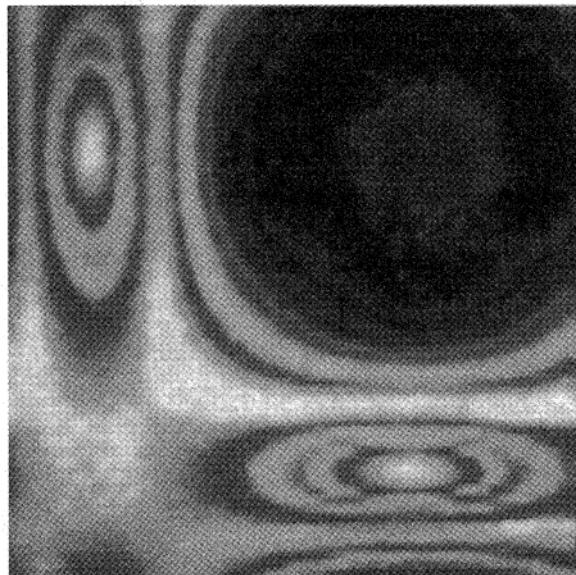
### Self-Organizing Structures

Now we would like to go over to the second part of our excursion into the realm of *Nonlinear Physics – Chaos and Order*, to the problem of order arising from disorder, to the problem of self-birth or self-generation of structures.

Figure 8.5 illustrates the birth of an ordered structure out of an initially disordered system as observed within the computer simulation of the two-dimensional medium depicted by the lattice model referred to in the context of Fig. 8.3. Out of disorder order is seen to arise – but maybe that happens only in computer experiments? Maybe, the in reality ubiquitous presence of noise and fluctuations does prevent the formation of such a self-organization in Nature?

A good example known from reality is the well known (see the book by the authors and/or the one by Haken) self-generation of a regular hexagonal lattice out of an initially mixed inhomogeneous liquid. This (Rayleigh-Bénard) lattice is the result of convective flows emerging in a horizontal layer of the liquid (in the given case it is a silicon oil) when uniformly heated from below. In this process the upper surface is free and surface tension plays an important role. Visualization is facilitated by adding aluminium powder to the liquid.

The hexagonal convective lattice arises only when the temperature difference between the heated lower liquid layer and the cold upper surface



**Fig. 8.5.** Spatial field distribution of the nonlinear lattice

exceeds a certain critical value. It is only in this case that the molecular heat exchange (in the absence of macroscopic motion in the respective liquid layer) is not able to comply with its “task” of transporting the temperature, thus leading to convective instabilities within the layer. They start off from fluctuations (read “disorder”!) and yield vortex motion with characteristic scales which depend on the thickness of the layer. (This instability is related to the fact that the lighter heated part of the liquid follows Archimedes’ law and moves upward forcing the heavier cold parts of the fluid to move downwards.) Thus we have again an instability. But now in the opposite sense: not giving rise to chaos but to order, to regular structures! The instability of a quiet, equilibrium state of the liquid leads to the preferred growth of the motion at individual characteristic scales at the expense of others and thus in the very end to regularly structured flows.

Such regular structured flows of liquids are quite often met in Nature. Amongst them are for example large vortices in the atmosphere of the Earth and also the regular structures of “topographical vortices” which we saw already in Fig. 8.1. These topographical vortices of volcanic material were formed as a consequence of the particular profile of the shelf.

But nevertheless, when we talk about the self-organization, self-generation of structures and follow their evolution towards completeness we (possibly subconsciously) have in mind something greater and by far more fundamental and unusual than simply the appearance of regular structures similar to the wave-like ones depicted in Fig. 8.6. This figure illustrates the regular three-



**Fig. 8.6.** Three-dimensional wavestructures [illustration courtesy of Ming-Yang-Su and M. van Dyke, Ed. *An Album of Fluid Motion* (The Parabolic Press, Stanford 1982) Fig. 194]

dimensional structure of waves on the surface of a fluid. Due to instability they arise from a homogeneous train of steep plane Stokes waves.

### Disorder and Structures

The notion of turbulence as a spatio-temporal chaos of structures is well-known to all of us. Just think about the evolution of smoke rings! The concept is also applicable to anisotropic media and in particular to shear flows. Structures arise as the result of evolving shear instabilities. Their dynamics and the character of their interaction along the flow becomes more complicated and as a result one observes eventually (downstream) the birth of turbulence.

To determine the general laws of the nonlinear dynamics of such structures it is convenient to use artificial nonequilibrium systems that can be prepared in the laboratory or even at home. They allow for detailed studies and a good understanding of the mentioned universality of the topology of these structures and their respective transformations in interactions with each other. One of the possible variants of such devices or artificial “media” can be set up at home with the help of an amateur video camera and a TV set: direct your camera towards the screen of your TV set and transmit the obtained signal of the camera into the video input of your TV set. This “camera + TV + feedback” system is self-excited for sufficiently strong amplification. Naturally, the amplification coefficient has to exceed the characteristic critical

value not just temporally but also spatially (which you may achieve by coming closer to the screen and accordingly stretching the picture).

As the result of such self-exitations you observe video pictures on the screen of the TV. In the direct sense they have nothing in common with the pattern “seen” by your video camera. If we experiment in a dark room then we find out that this system is characterized by “hard” excitation. In the dark it is necessary to generate a light pulse to “ignite” the screen. You can watch how the light of the flame of a lighted match brought within the aperture of the objective spreads out across the screen (according to the laws of nonlinear diffusion). The pictures which we thus obtain resemble the ones of the fractal structures looked at earlier. We observe spirals, “viscous fingers” and many others on the screen of our monitor, see the book by Gaponov-Grekhov and Rabinovich.

The dynamics of the structures in our home experiment with unstable media are of an amazing variety. Depending on the particular conditions, they may either be chaotic or regular. With the help of a separate camera directed towards the screen we may obtain pictures showing self-oscillations of our “camera + TV + feedback” system: The video structures replace each other periodically as time progresses.

### Multistability and Memory

We may also use the “camera + TV + feedback” system for completely unexpected purposes, say, for obtaining parallel backward information. Let us ponder a bit about this point.

As many as three decades ago it was clarified that a good number of the processes associated with the vision of a frog (including practically all computations related to the recognition of simple objects like insects) takes place in the retina of its eye. In the language of our computerized world the retina is just an ensemble of parallel processors which spend almost no time in communication with each other. Why are they not communicating with each other? – In the case of a network of only twelve communicating processors almost *all* time is needed for the exchange of information so that there is essentially no time left for the actual evaluation of the information, that is for the recognition of the object.

The simplest architecture of a computer with parallel information processing of pictures is represented by a rectangular matrix of processors, each communicating only with its neighbour. Onto such a matrix we may by the aid of a video camera project the picture which is to be evaluated. Devices of this type are called *Cellular Logic Image Processors* or *Automata*. If we assume that each element of the matrix can only perform a small number of operations, then the number of states which it can receive as a function of the amplitude of the incident signal is also small. Let us try to employ as such a *Cellular Processor* the screen of the colour TV set in our TV-system with feedback. By the aid of a lattice of “open boxes” arranged in front of

the screen we divide its area into small cells (the dimensions of which should naturally not be smaller than the ones of the smallest points our TV set can reproduce stably). Each of the cells can now be in one of four independent states (the nonexcited ground state versus three excited states represented by red, blue and green). Thus the number of stably existing pictures (attractors) in our “camera + TV + feedback” system may be sufficiently large. The system may be multistable. The stability of the pictures allows us on the one hand to remember them, on the other hand to distinguish different ones. Some nice illustrations of the phenomenon of multistability for our TV-system are shown in our book.

### Nonlinear Dynamics in Society

We would like to add that models and methods of nonlinear dynamics themselves are also continuously developing. Let us mention just one comparatively new direction, the modelling of nonlinear dynamics of extended systems with the help of cellular automata. In this case the “medium” is treated as a set of discrete cells with given rules for their interaction with each other and evolving in discrete time. Such models turned out to be extremely effective for traditional problems of physics as well as, in particular, for problems related to the modelling of evolution processes and of self-teaching (neural networks).

Already for extremely simple interaction laws between each other, the life of cellular automata may turn out to be highly sophisticated and even chaotic as demonstrated in our book. After the experience gained so far with such nonlinear systems we should no longer be surprised by this. If we follow the time evolution of complicated non-repeated spatial structures in cellular automata we will see that neighbouring cells excite each other while the ones which are not so close to each other slow the process down. The colour indicates the intensity of the excitation; black corresponds to no excitation at all.

In spite of our bad experience with predictions we dare to assert that within the coming decade the methods of nonlinear physics and nonlinear dynamics will come into vogue not only for physicians and ecologists but also for economists, sociologists and geographers.

Indeed, as clarified most recently, the growth of cities and the evolution of the network of urban and intercity transport resemble strongly the growth of fractal clusters in models with “limited” diffusion.

The majority of specialists concur nowadays that the idea of dynamic equilibrium which lies at the foundation of traditional economical models is not satisfactory. It is realized that it will inevitably be necessary to build nonlinear dynamic models. One approach as proposed by Bak, Tang and Wiesenfeld is named *Self-Organized Criticality*. When the modelled system reaches a critical state, it experiences a self-reorganization and acquires a new form which may be predicted with a sufficient degree of probability. A similar

type of prediction – admittedly applied to the past! – was for example made with the aid of the method of critical dynamics in models of arms races between countries which subsequently entered into open war with each other, see the paper by A. Superstein. The methods of critical dynamics are indeed universal. This universality manifests itself also in the character of the changes observed in the dynamics of completely different processes due to changes in the parameters determining the evolution of these processes. Thus it is, e.g. that the same changes (bifurcations) are experienced within the mentioned arms races and the oscillations observed in experiments with microwave generators: a bifurcation (period) doubling and the transition to chaos.

## Further Reading

- A. Superstein: Nature **309**, 303 (1984)
- P. Bak, C. Tang, K. Wiesenfeld: Self-Organized Criticality, Phys. Rev. Lett. **59**, 381 (1987)
- A.V. Gaponov-Grekhov, M.I. Rabinovich: *Nonlinearities in Action* (Springer, Berlin, Heidelberg, New York 1992)
- H. Haken: *Synergetics* (Springer, Berlin, Heidelberg, New York 1978)
- M.V. Nezlin, E.N. Slezkin: *Rossby Vortices, Spiral Structures, Solitons* (Springer, Berlin, Heidelberg, New York 1992)
- H.-O. Peitgen, P.H. Richter: *The Beauty of Fractals* (Springer, Berlin, Heidelberg, New York 1992)

## 8.2 Quantum Chaos<sup>2</sup>

By Martin C. Gutzwiller<sup>3</sup>

*Does chaos lurk in the smooth, wavelike quantum world? Recent work shows that the answer is yes – symptoms of chaos enter even into the wave patterns associated with atomic energy levels*

In 1917 Albert Einstein wrote a paper that was completely ignored for 40 years. In it he raised a question that physicists have only recently begun asking themselves: What would classical chaos, which lurks everywhere in our world, do to quantum mechanics, the theory describing the atomic and subatomic worlds? The effects of classical chaos, of course, have long been observed – Kepler knew about the irregular motion of the moon around the earth, and Newton complained bitterly about the phenomenon. At the end of the 19th century, the American astronomer George William Hill demonstrated that the irregularity is the result entirely of the gravitational pull of the sun. Shortly thereafter, the great French mathematician-astronomer-physicist Henri Poincaré surmised that the moon's motion is only a mild case of a congenital disease affecting nearly everything. In the long run, Poincaré realized, most dynamic systems show no discernible regularity or repetitive pattern. The behavior of even a simple system can depend so sensitively on its initial conditions that the final outcome is uncertain.

At about the time of Poincaré's seminal work on classical chaos, Max Planck started another revolution, which would lead to the modern theory of quantum mechanics. The simple systems that Newton had studied were investigated again, but this time on the atomic scale. The quantum analogue of the humble pendulum is the laser; the flying cannonballs of the atomic world consist of beams of protons or electrons, and the rotating wheel is the spinning electron (the basis of magnetic tapes). Even the solar system itself is mirrored in each of the atoms found in the periodic table of the elements.

---

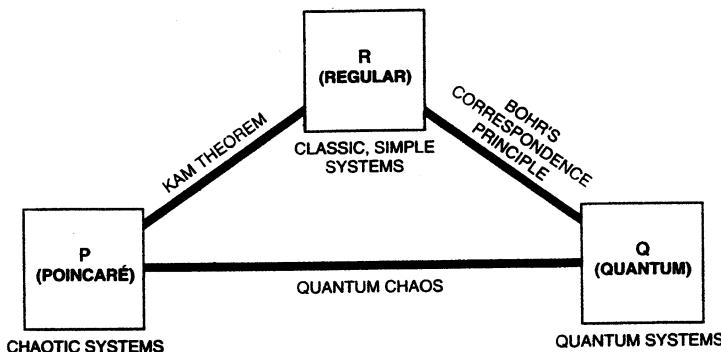
<sup>2</sup> The original appeared in *Scientific American*, January 1992. Reprinted with permission. ©(1991) by Scientific American, Inc. All rights reserved.

<sup>3</sup> *Martin C. Gutzwiller* is a member of the US National Academy of Sciences. He is at the research staff at the IBM Thomas J. Watson Research Center in Yorktown Heights, N.Y. and also serves as adjunct professor of metallurgy at the Columbia University School of Engineering. Born in Switzerland, he received his education in Swiss public schools and at the Federal Institute of Technology in Zurich, where he obtained a degree in physics and mathematics in 1950. His Ph.D. in physics was awarded by the University of Kansas in 1953. Gutzwiller joined IBM in 1960 after working for seven years in the Exploration and Production Research Laboratory of Shell Oil Company. At IBM his focus was first on the interaction of electrons in metals and is now on the relations between classical and quantum mechanics. He is also interested in celestial mechanics and the history of astronomy and physics. In addition, he is an amateur musician, a collector of old science books and a hiker.

Perhaps the single most outstanding feature of the quantum world is its smooth and wavelike nature. This feature leads to the question of how chaos makes itself felt when moving from the classical world to the quantum world. How can the extremely irregular character of classical chaos be reconciled with the smooth and wavelike nature of phenomena on the atomic scale? Does chaos exist in the quantum world?

Preliminary work seems to show that it does. Chaos is found in the distribution of energy levels of certain atomic systems; it even appears to sneak into the wave patterns associated with those levels. Chaos is also found when electrons scatter from small molecules. I must emphasize, however, that the term “quantum chaos” serves more to describe a conundrum than to define a well-posed problem.

Considering the following interpretation of the bigger picture may be helpful in coming to grips with quantum chaos. All our theoretical discussions of mechanics can be somewhat artificially divided into three compartments (see Fig. 8.7) although nature recognizes none of these divisions.



**Fig. 8.7.** Mechanics is traditionally (and artificially) divided into the three compartments depicted here, which are linked together by several connections. Quantum chaos is concerned with establishing the relation between boxes *P* and *Q*

Elementary classical mechanics falls in the first compartment. This box contains all the nice, clean systems exhibiting simple and regular behavior, and so I shall call it *R*, for regular. Also contained in *R* is an elaborate mathematical tool called perturbation theory, which is used to calculate the effects of small interactions and extraneous disturbances, such as the influence of the sun on the moon’s motion around the earth. With the help of perturbation theory, a large part of physics is understood nowadays as making relatively mild modifications of regular systems. Reality, though, is much more complicated; chaotic systems lie outside the range of perturbation theory, and they constitute the second compartment.

Since the first detailed analyses of the systems of the second compartment were done by Poincaré, I shall call this box  $P$  in his honor. It is stuffed with the chaotic dynamic systems that are the bread and butter of science [see Crutchfield et al.]. Among these systems are all the fundamental problems of mechanics, starting with three, rather than only two, bodies interacting with one another, such as the earth, moon and sun, or the three atoms in the water molecule, or the three quarks in the proton.

Quantum mechanics, as it has been practiced for about 90 years, belongs in the third compartment, called  $Q$ . After the pioneering work of Planck, Einstein and Niels Bohr, quantum mechanics was given its definitive form in four short years, starting in 1924. The seminal work of Louis de Broglie, Werner Heisenberg, Erwin Schrödinger, Max Born, Wolfgang Pauli and Paul Dirac has stood the test of the laboratory without the slightest lapse. Miraculously, it provides physics with a mathematical framework that, according to Dirac, has yielded a deep understanding of “most of physics and all of chemistry”. Nevertheless, even though most physicists and chemists have learned how to solve special problems in quantum mechanics, they have yet to come to terms with the incredible subtleties of the field. These subtleties are quite separate from the difficult, conceptual issues having to do with the interpretation of quantum mechanics.

The three boxes  $R$  (classic, simple systems),  $P$  (classic chaotic systems) and  $Q$  (quantum systems) are linked by several connections. The connection between  $R$  and  $Q$  is known as Bohr’s correspondence principle. The correspondence principle claims, quite reasonably, that classical mechanics must be contained in quantum mechanics in the limit where objects become much larger than the size of atoms. The main connection between  $R$  and  $P$  is the Kolmogorov-Arnold-Moser (KAM) theorem. The KAM theorem provides a powerful tool for calculating how much of the structure of a regular system survives when a small perturbation is introduced, and the theorem can thus identify perturbations that will cause a regular system to undergo chaotic behavior.

Quantum chaos is concerned with establishing the relation between boxes  $P$  (chaotic systems) and  $Q$  (quantum systems). In establishing this relation, it is useful to introduce a concept called phase space. Quite amazingly, this concept, which is now so widely exploited by experts in the field of dynamic systems, dates back to Newton.

The notion of phase space can be found in Newton’s *Mathematical Principles of Natural Philosophy*, published in 1687. In the second definition of the first chapter, entitled “Definitions”, Newton states (as translated from the original Latin in 1729): “The quantity of motion is the measure of the same, arising from the velocity and quantity of matter conjointly.” In modern English, this means that for every object there is a quantity, called momentum, which is the product of the mass and velocity of the object.

Newton gives his laws of motion in the second chapter, entitled “Axioms, or Laws of Motion.” The second law says that the change of motion is proportional to the motive force impressed. Newton relates the force to the change of momentum (not to the acceleration, as most textbooks do).

Momentum is actually one of two quantities that, taken together, yield the complete information about a dynamic system at any instant. The other quantity is simply position, which determines the strength and direction of the force. Newton’s insight into the dual nature of momentum and position was put on firmer ground some 150 years later by two mathematicians, William Rowan Hamilton and Karl Gustav Jacob Jacobi. The pairing of momentum and position is no longer viewed in the good old Euclidean space of three dimensions; instead it is viewed in phase space, which has six dimensions, three dimensions for position and three for momentum.

The introduction of phase space was a wonderful step from a mathematical point of view, but it represents a serious setback from the standpoint of human intuition. Who can visualize six dimensions? In some cases, fortunately, the phase space can be reduced to three or, even better, two dimensions.

Such a reduction is possible in examining the behavior of a hydrogen atom in a strong magnetic field. The hydrogen atom has long been a highly desirable system because of its simplicity: a lone electron moves around a lone proton. And yet the classical motion of the electron becomes chaotic when the magnetic field is turned on. How can we claim to understand physics if we cannot explain this basic problem?

Under normal conditions, the electron of a hydrogen atom is tightly bound to the proton. The behavior of the atom is governed by quantum mechanics. The atom is not free to take on any arbitrary energy; it can take on only discrete, or quantized, energies. At low energies, the allowed values are spread relatively far apart. As the energy of the atom is increased, the atom grows bigger, because the electron moves farther from the proton, and the allowed energies get closer together. At high enough energies (but not too high, or the atom will be stripped of its electron!), the allowed energies get very close together into what is effectively a continuum, and it now becomes fair to apply the rules of classical mechanics.

Such a highly excited atom is called a Rydberg atom [see Kleppner et al.] Rydberg atoms inhabit the middle ground between the quantum and the classical worlds, and they are therefore ideal candidates for exploring Bohr’s correspondence principle, which connects boxes  $Q$  (quantum phenomena) and  $R$  (classic phenomena). If a Rydberg atom could be made to exhibit chaotic behavior in the classical sense, it might provide a clue as to the nature of quantum chaos and thereby shed light on the middle ground between boxes  $Q$  and  $P$  (chaotic phenomena).

A Rydberg atom exhibits chaotic behavior in a strong magnetic field, but to see this behavior we must reduce the dimension of the phase space. The first step is to note that the applied magnetic field defines an axis of symmetry

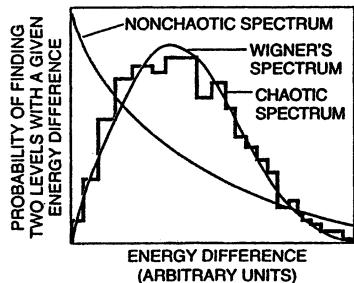
through the atom. The motion of the electron takes place effectively in a two-dimensional plane, and the motion around the axis can be separated out; only the distances along the axis and from the axis matter. The symmetry of motion reduces the dimension of the phase space from six to four.

Additional help comes from the fact that no outside force does any work on the electron. As a consequence, the total energy does not change with time. By focusing attention on a particular value of the energy, one can take a three-dimensional slice – called an energy shell – out of the four-dimensional phase space. The energy shell allows one to watch the twists and turns of the electron, and one can actually see something resembling a tangled wire sculpture. The resulting picture can be simplified even further through a simple idea that occurred to Poincaré. He suggested taking a fixed two-dimensional plane (called a Poincaré section, or a surface of section) through the energy shell and watching the points at which the trajectory intersects the surface. The Poincaré section reduces the tangled wire sculpture to a sequence of points in an ordinary plane.

A Poincaré section for a highly excited hydrogen atom in a strong magnetic field shows regions of the phase space where the points are badly scattered, indicating chaotic behavior. Such scattering is a clear symptom of classical chaos, and it allows one to separate systems into either box  $P$  or box  $R$ .

What does the Rydberg atom reveal about the relation between boxes  $P$  and  $Q$ ? I have mentioned that one of the trademarks of a quantum mechanical system is its quantized energy levels, and in fact the energy levels are the first place to look for quantum chaos. Chaos does not make itself felt at any particular energy level, however; rather its presence is seen in the spectrum, or distribution, of the levels. Perhaps somewhat paradoxically, in a nonchaotic quantum system the energy levels are distributed randomly and without correlation, whereas the energy levels of a chaotic quantum system exhibit strong correlations, see Fig. 8.8. The levels of the regular system are often close to one another, because a regular system is composed of smaller subsystems that are completely decoupled. The energy levels of the chaotic system, however, almost seem to be aware of one another and try to keep a safe distance. A chaotic system cannot be decomposed; the motion along one coordinate axis is always coupled to what happens along the other axis.

The spectrum of a chaotic quantum system was first suggested by Eugene P. Wigner, another early master of quantum mechanics. Wigner observed, as had many others, that nuclear physics does not possess the safe underpinnings of atomic and molecular physics; the origin of the nuclear force is still not clearly understood. He therefore asked whether the statistical properties of nuclear spectra could be derived from the assumption that many parameters in the problem have definite, but unknown, values. This rather vague starting point allowed him to find the most probable formula for the distribution. Oriol Bohigas and Marie-Joya Giannoni of the Institute of Nuclear Physics



**Fig. 8.8.** The Energy Spectrum, or distribution of energy levels, differs markedly between chaotic and nonchaotic quantum systems. For a nonchaotic system, such as a molecular hydrogen ion ( $H_2^+$ ), the probability of finding two energy levels close to each other is quite high. In the case of a chaotic system such as a Rydberg atom in a strong magnetic field, the probability is low. The chaotic spectrum closely matches the typical nuclear spectrum derived many years ago by Eugene P. Wigner

in Orsay, France, first pointed out that Wigner's distribution happens to be exactly what is found for the spectrum of a chaotic dynamic system.

Chaos does not seem to limit itself to the distribution of quantum energy levels, however; it even appears to work its way into the wavelike nature of the quantum world. The position of the electron in the hydrogen atom is described by a wave pattern. The electron cannot be pinpointed in space; it is a cloudlike smear hovering near the proton. Associated with each allowed energy level is a stationary state, which is a wave pattern that does not change with time. A stationary state corresponds quite closely to the vibrational pattern of a membrane that is stretched over a rigid frame, such as a drum.

The stationary states of a chaotic system have surprisingly interesting structure, as demonstrated in the early 1980s by Eric Heller of the University of Washington. He and his students calculated a series of stationary states for a two-dimensional cavity in the shape of a stadium. The corresponding problem in classical mechanics was known to be chaotic, for a typical trajectory quickly covers most of the available ground quite evenly. Such behavior suggests that the stationary states might also look random, as if they had been designed without rhyme or reason. In contrast, Heller discovered that most stationary states are concentrated around narrow channels that form simple shapes inside the stadium, and he called these channels "scars". Similar structure can also be found in the stationary states of a hydrogen atom in a strong magnetic field. In the original paper in the *Scientific American* you find at this place a nice illustration due to Heller and an impressive colored illustration of *Stationary states*, or wave patterns, associated with the energy levels of a Rydberg atom (a highly excited hydrogen atom) in a strong magnetic field which can exhibit chaotic qualities. The smoothness of the quantum wave forms is preserved from point to point, but when one steps back to view the whole picture, the fingerprint of chaos emerges.

It is possible to connect the chaotic signature of the energy spectrum to ordinary classical mechanics. A clue to the prescription is provided in Einstein's 1917 paper. He examined the phase space of a regular system from box  $R$  and described it geometrically as filled with surfaces in the shape of a donut; the motion of the system corresponds to the trajectory of a point over the surface of a particular donut. The trajectory winds its way around the surface of the donut in a regular manner, but it does not necessarily close on itself.

In Einstein's picture, the application of Bohr's correspondence principle to find the energy levels of the analogous quantum mechanical system is simple. The only trajectories that can occur in nature are those in which the cross section of the donut encloses an area equal to an integral multiple of Planck's constant,  $\hbar$  ( $2\pi$  times the fundamental quantum of angular momentum, having the units of momentum multiplied by length). It turns out that the integral multiple is precisely the number that specifies the corresponding energy level in the quantum system.

Unfortunately, as Einstein clearly saw, his method cannot be applied if the system is chaotic, for the trajectory does not lie on a donut, and there is no natural area to enclose an integral multiple of Planck's constant. A new approach must be sought to explain the distribution of quantum mechanical energy levels in terms of the chaotic orbits of classical mechanics.

Which features of the trajectory of classical mechanics help us to understand quantum chaos? Hill's discussion of the moon's irregular orbit because of the presence of the sun provides a clue. His work represented the first instance where a particular periodic orbit is found to be at the bottom of a difficult mechanical problem. (A periodic orbit is like a closed track on which the system is made to run; there are many of them, although they are isolated and unstable.) Inspiration can also be drawn from Poincaré, who emphasized the general importance of periodic orbits. In the beginning of this three-volume work, *The New Methods of Celestial Mechanics*, which appeared in 1892, he expresses the belief that periodic orbits "offer the only opening through which we might penetrate into the fortress that has the reputation of being impregnable". Phase space for a chaotic system can be organized, at least partially, around periodic orbits, even though they are sometimes quite difficult to find.

In 1970 I discovered a very general way to extract information about the quantum mechanical spectrum from a complete enumeration of the classical periodic orbits. The mathematics of the approach is too difficult to delve into here, but the main result of the method is a relatively simple expression called a trace formula. The approach has now been used by a number of investigators, including Michael V. Berry of the University of Bristol, who has used the formula to derive the statistical properties of the spectrum.

I have applied the trace formula to compute the lowest two dozen energy levels for an electron in a semiconductor lattice, near one of the care-

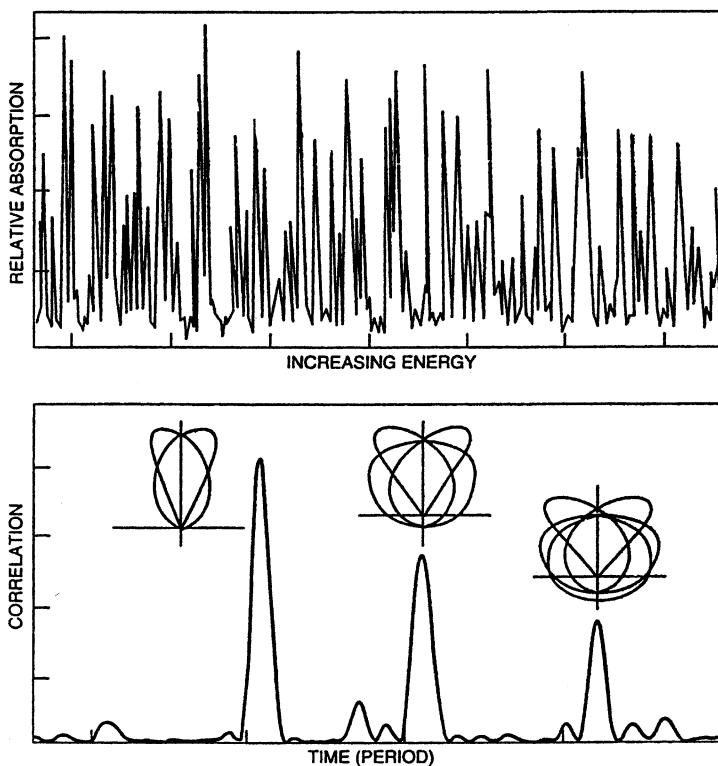
fully controlled impurities. (The semiconductor, of course, is the basis of the marvelous devices on which modern life depends; because of its impurities, the electrical conductivity of the material is halfway between that of an insulator, such as plastic, and that of a conductor, such as copper.) The trajectory of the electron can be uniquely characterized by a string of symbols, which has a straightforward interpretation. The string is produced by defining an axis through the semiconductor and simply noting when the trajectory crosses the axis. A crossing to the “positive” side of the axis gets the symbol +, and a crossing to the “negative” side gets the symbol −.

A trajectory then looks exactly like the record of a coin toss. Even if the past is known in all detail – even if all the crossings have been recorded – the future is still wide open. The sequence of crossings can be chosen arbitrarily. Now, a periodic orbit consists of a binary sequence that repeats itself; the simplest such sequence is (+−), the next is (++−), and so on. (Two crossings in a row having the same sign indicate that the electron has been trapped temporarily.) All periodic orbits are thereby enumerated, and it is possible to calculate an approximate spectrum with the help of the trace formula. In other words, the quantum mechanical energy levels are obtained in an approximation that relies on quantities from classical mechanics only.

The classical periodic orbits and the quantum mechanical spectrum are closely bound together through the mathematical process called Fourier analysis [see *Bracewell*]. The hidden regularities in one set, and the frequency with which they show up, are exactly given by the other set. This idea was used by John B. Delos of the College of William and Mary and Dieter Wintgen of the Max Planck Institute for Nuclear Physics in Heidelberg to interpret the spectrum of the hydrogen atom in a strong magnetic field.

Experimental work on such spectra has been done by Karl H. Welge and his colleagues at the University of Bielefeld, who have excited hydrogen atoms nearly to the point of ionization, where the electron tears itself free of the proton. The energies at which the atoms absorb radiation appear to be quite random (see upper part of Fig. 8.9), but a Fourier analysis converts the jumble of peaks into a set of well-separated peaks (see lower part of Fig. 8.9). The important feature here is that each of the well-separated peaks corresponds precisely to one of several standard classical periodic orbits. Poincaré’s insistence on the importance of periodic orbits now takes on a new meaning. Not only does the classical organization of phase space depend critically on the classical periodic orbits, but so too does the understanding of a chaotic quantum spectrum.

So far I have talked only about quantum systems in which an electron is trapped or spatially confined. Chaotic effects are also present in atomic systems where an electron can roam freely, as it does when it is scattered from the atoms in a molecule. Here energy is no longer quantized, and the



**Fig. 8.9.** Absorption of light by a hydrogen atom in a strong magnetic field appears to vary randomly as a function of energy (top), but when the data are analyzed according to the mathematical procedure called Fourier analysis, a distinct pattern emerges (bottom). Each peak in the bottom panels has associated with it a specific classical periodic orbit (figures next to peaks)

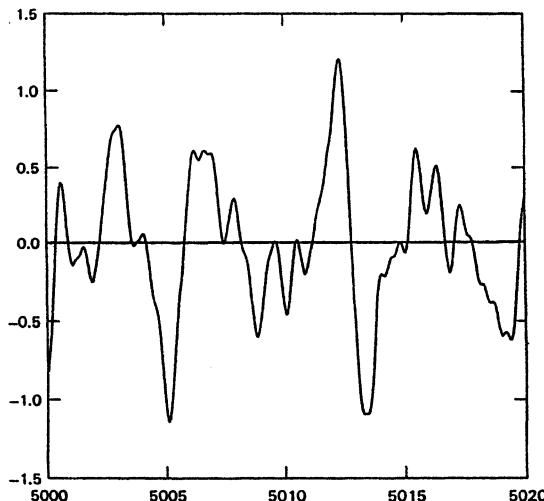
electron can take on any value, but the effectiveness of the scattering depends on the energy.

Chaos shows up in quantum scattering as variations in the amount of time the electron is temporarily caught inside the molecule during the scattering process. For simplicity, the problem can be examined in two dimensions. To the electron, a molecule consisting of four atoms looks like a small maze. When the electron approaches one of the atoms, it has two choices: it can turn left or right. Each possible trajectory of the electron through the molecule can be recorded as a series of left and right turns around the atoms, until the particle finally emerges. All of the trajectories are unstable: even a minute change in the energy or the initial direction of the approach will cause a large change in the direction in which the electron eventually leaves the molecule.

The chaos in the scattering process comes from the fact that the number of possible trajectories increases rapidly with path length. Only an interpre-

tation from the quantum mechanical point of view gives reasonable results; a purely classical calculation yields nonsensical results. In quantum mechanics, each classical trajectory of the electron is used to define a little wavelet that winds its way through the molecule. The quantum mechanical result follows from simply adding up all such wavelets.

Recently I have done a calculation of the scattering process for a special case in which the sum of the wavelets is exact. An electron of known momentum hits a molecule and emerges with the same momentum. The arrival time for the electron to reach a fixed monitoring station varies as a function of the momentum, and the way in which it varies is what is so fascinating about this problem. The arrival time fluctuates smoothly over small changes in the momentum, but over large changes a chaotic imprint emerges, which never settles down to any simple pattern (see right part of Fig. 8.10).



**Fig. 8.10.** Chaotic variation (characterizes the time it takes for a scattered electron of transiting a molecule) of known momentum to reach a fixed monitoring station. Arrival time varies as a function of the electron's momentum. The variation is smooth when changes in the momentum are small but exhibits a complex chaotic pattern when the changes are large. The quantity shown on the vertical axis, the phase shift, is a measure of the time delay

A particularly tantalizing aspect of the chaotic scattering process is that it may connect the mysteries of quantum chaos with the mysteries of number theory. The calculation of the time delay leads straight into what is probably the most enigmatic object in mathematics, Riemann's zeta function. Actually, it was first employed by Leonhard Euler in the middle of the 18th century to show the existence of an infinite number of prime numbers (integers that cannot be divided by any smaller integer other than one). About a century

later Bernhard Riemann, one of the founders of modern mathematics, employed the function to delve into the distribution of the primes. In his only paper on the subject, he called the function by the Greek letter zeta.

The zeta function is a function of two variables,  $x$  and  $y$  (which exist in the complex plane). To understand the distribution of prime numbers, Riemann needed to know when the zeta function has the value of zero. Without giving a valid argument, he stated that it is zero only when  $x$  is set equal to  $\frac{1}{2}$ . Vast calculations have shown that he was right without exception for the first billion zeros, but no mathematician has come even close to providing a proof. If Riemann's conjecture is correct, all kinds of interesting properties of prime numbers could be proved.

The values of  $y$  for which the zeta function is zero form a set of numbers that is much like the spectrum of energies of an atom. Just as one can study the distribution of energy levels in the spectrum, so can one study the distribution of zeros for the zeta function. Here the prime numbers play the same role as the classical closed orbits of the hydrogen atom in magnetic field: the primes indicate some of the hidden correlations among the zeros of the zeta function.

In the scattering problem the zeros of the zeta function give the values of the momentum where the time delay changes strongly. The chaos of the Riemann zeta function is apparent in a theorem that has only recently been proved: the zeta function fits locally any smooth function. The theorem suggests that the function may describe all the chaotic behavior a quantum system can exhibit. If the mathematics of quantum mechanics could be handled more skillfully, many examples of locally smooth, yet globally chaotic, phenomena might be found.

## Further Reading

M.V. Berry, I.C. Percival, N.O. Weiss (Eds.): Dynamical Chaos in *Proc. Royal Soc. (London)*, **A413** No. 1844, pp. 1–199; September 8, 1987

R.N. Bracewell: “The Fourier Transform” in *Scientific American*, June 1989

J.P. Crutchfield, J.D. Farmer, N.H. Packard, R.S. Shaw: “Chaos” in *Scientific American*, December 1986

H. Friedrich, D. Wintgen: The Hydrogen Atom in a Uniform Magnetic Field: An Example of Chaos. *Physics Reports* **183**, No. 2, pp. 37–79; November 1989

M.C. Gutzwiller: *Chaos in Classical and Quantum Mechanics*. (Springer-Verlag, Berlin, Heidelberg, New York 1990)

D. Kleppner, M.G. Littman, M.L. Zimmerman: “Highly Excited Atoms” in *Scientific American*, May 1981

H.J. Korch, H-J. Jodl: *Chaos, a Program Collection for the PC* (Springer-Verlag, Berlin Heidelberg 1994)

T. Uzer, D. Farrelly, J.A. Milligen, P.E. Raines, J.P. Skelton: “Celestial Mechanics on a Microscopic Scale”. *Science*, **253**, pp. 42–48; July 5, 1991

A. Voros, M.-J. Giannoni, J. Zinn-Justin (Eds.): *Chaos and Quantum Physics*. (Elsevier, Amsterdam 1991)

# Hints for Solution

## Chapter 1

- 1.1  $1.3.5 \dots (2n - 1) / [(2n)! 2^n a^n]$  simplifies further.
- 1.2 Notice the word “some”. This procedure greatly increases the number of good functions.
- 1.3 Use  $(a - ib)$ , the complex conjugate of  $(a + ib)$ .
- 1.5  $\cos 3\theta = \cos^3 \theta - 3 \cos \theta \sin^2 \theta \sin 3\theta = 3 \cos^2 \theta \sin \theta - \sin^3 \theta$ .
- 1.6 a)  $\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} f'_n(x) F(x) dx = - \lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} f_n(x) F'(x) dx$ .  
b) From (1.21) the answer is  $-1$ .
- 1.7 Consider both positive and negative values of  $|k|$  together.
- 1.8 
$$\begin{aligned} E(t) &= 1/\sqrt{2\pi} \int_{-\infty}^{\infty} A_{\omega} e^{i\omega t} d\omega \\ &= \sqrt{2\pi} \int_{-\infty}^{\infty} A(f) e^{2\pi i f t} df \end{aligned}$$

$10^8 + 75 \times 10^3 \gtrsim f \gtrsim 10^8 - 75 \times 10^3$  Hz if the radio is turned to 100 MHz

1.9 a)  $A(k) = A^*(-k)$   
b)  $|A_k| \cos(\theta_k + k_n) = |A_k| (\cos \theta_k \cos kx - \sin \theta_k \sin kx)$   
c) Four parameters are available for every  $k > 0$ .

1.10 a)  $f_e(x) = f(x) + f(-x)/2$ .  
b)  $C(k) = [A(k) + A(-k)]/\sqrt{2}$ .  
d) Refer to (1.16) and (1.32).  
e) Do  $1/(2\sqrt{\pi}) \cos kx, 1/(2\sqrt{\pi}) \sin kx$ , form a complete orthonormal set?

1.11 a)  $A(k) = (1/\sqrt{2\pi}) 1/(a + ik)$ .  
b)  $F(x)$  is nonsymmetric but real.

1.12 b) Note the reciprocal relation which says that if  $A(k)$  is the FT of  $F(x)$ , (1.26), then  $F(k)$  is the FT of  $A(-x)$ , [(1.33) with  $k' \rightarrow -x, x \rightarrow k$ ].  
c) Consider “smudged” (why?) square waves, and note (a) and (b).  
d)  $\sqrt{n/\pi} e^{-nx^2}, (\sin nx)/(\pi x)$ .

- 1.13 a) The FT of  $(1/\sqrt{2\pi})e^{-|x|/n}$  is  $1/(\pi n)(n^{-2} + k^2)^{-1}$ .  
 b) See Problem 1.3.  
 c) Expand  $F(x)$  in a Taylor series about  $x = 0$ .  
 d) Here the  $a_n(k)$  are good functions differentiable at  $k = 0$  and equal to the generalized function 1 (why?). To find the  $a_n(k)$ , differentiate the desired transform with respect to  $1/n^2$  in only the numerator! Split your derivative into two terms with integrating factors  $[1/2 + kn^2/4]$ ,  $[1/2 - k/(4/n^2)]$ , respectively.
- 1.14 Both sides can be written as the double integral

$$1/\sqrt{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A_1(k)e^{-ikx} F_2(x) dk dx .$$

- 1.15 a) See Problem 1.14.  
 b)  $F(x) = 1/2\pi \int_{-\infty}^{\infty} [G(k)/B(k)] e^{ikx} dk .$
- 1.16 See Problem 1.9a.
- 1.17 a)  $A(k) = 1/\sqrt{2\pi} e^{-ikx_0}$ , or  $1/\sqrt{2\pi} e^{-ikx_0 - n(x-x_0)^2}$   
 if  $\varrho(0, x) = g_n(x - x_0)$ .  
 b) Integrate over  $k$ . Refer to (1.18).
- 1.18 c) The factor assures that  $f_n(x)$  is “good” at infinity. See *Lighthill* (Further Reading) p.22 for a proof that (1.22) is satisfied.
- 1.19 Consider

$$\lim_{n \rightarrow \infty} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} A(k) f_n(k) e^{ikx} dk ,$$

where

$$A(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(x') e^{-ikx'} dx'$$

- with  $f_n(k)$  a sequence defining the generalized function unity, for example,  $f_n(k) = \exp(-k^2/4n)$ .
- 1.20 (iii) With  $u_i(x)$  real

$$\widetilde{A}_i \cong \int_{-N\Delta/2}^{N\Delta/2} u_i(x) F_A(x) dx .$$

**Chapter 2**

- 2.4 a)  $V_c = -IR$ .  
 b) Can you show that to first order in  $\alpha$

$$\frac{q_0}{C} e^{-\alpha t} \cos \omega_f t = \omega_f R q_0 e^{-\alpha t} \sin \omega_f t$$

at the point of inflection for  $q(t)$ ?

- 2.5 This is a general initial value problem. Try it. Alternatively show that  $Q = Q_0 + \int_0^t I dt$  for the given current.  
 2.6 a)  $(\alpha^2 - \omega_0^2)^{1/2} \simeq \alpha - \omega_0^2/2\alpha$   
 b)  $\phi \rightarrow -\pi/2$ ,  $|x| \rightarrow F_0/k$ .  
 2.7 b)  $\sim 106$   
 2.8 a)  $1/T \int_0^T \sin(\omega t) \cos(\omega t) dt = 0$   
 2.10  $\chi_2(\omega) = \left( \frac{2\alpha\omega}{\sqrt{2\pi m}} \right) \frac{1}{(\omega^2 - \omega_0^2)^2 + (2\alpha\omega)^2}$   
 measures the in-phase component of the velocity.

**Chapter 3**

3.1  $y(x, t) = 1/2 \left[ y_0(x - st) + y_0(x + st) + (1/s) \int_{x-st}^{x+st} v_0(x') dx' \right]$ .

- 3.6 b)  $(1 + \delta)^{1/2} \approx 1 + \delta/2$  for any  $\delta \ll 1$ .  
 3.7 a) See the derivation of the wave equation for a transverse force.  
 3.9 For nontrivial solutions, the determinant

$$\begin{bmatrix} w^2 - \alpha & \beta \\ \beta & w^2 - \alpha \end{bmatrix} \quad \alpha = (k_1 + k_2)/m \quad \beta = k_2/m$$

must vanish. The normal modes are

$$X_I = \sqrt{2/3} \sin(\pi I/3) \cos(t\sqrt{\alpha - \beta}) \quad I = 1, 2$$

$$X_I = \sqrt{2/3} \sin(2\pi I/3) \cos(t\sqrt{\alpha + \beta}) \quad I = 1, 2$$

i.e., the two masses move together (low frequency mode), or they move against each other (high frequency).

- b) See Example 6.4.  
 3.12 b) Consider
- $$\frac{1}{L} \int_{-L/2}^{+L/2} \exp \left[ \left( \frac{2\pi i}{L} \right) (n - m)x \right] dx$$
- c) The result is the integral (as opposed to series) Fourier transform of Chap. 1 for functions that may be Fourier represented.  
 3.13 a) On a unit cross-sectional area at  $x + \Delta x$  the force is  $p(x + \Delta x, t)$ .  
 b) Start by examining the solution  $L = \lambda/4$ .

- 3.14 b) The transverse force exerted on the end of the string is  $-T \frac{\partial y}{\partial x} \Big|_{x=0}$ .  
e)  $\frac{\partial y(x, t)}{\partial t} \rightarrow \frac{F_0 s}{T} \cos(\omega_0 t - \omega_0 x/s)$ .  
f) The analytic result corresponds to a pulse of width  $\sim s\sqrt{n}$  centered at  $st$ . Explain physically the origin of your result. Is there any way of obtaining the idealized result of part (e) from your result here?
- 3.15 a) Bear in mind that the boundary conditions must be satisfied throughout time.  
b) At resonance, for the electric analogue  $Z_{\text{parallel}} = \infty$   $Z_{\text{series}} = 0$ .
- 3.16 a) First see Problem 2.9(c).

$$z(x, t) = A \sum_n u_n(x) \frac{\omega_0^2}{\omega_n^2 - \omega_0^2} \left( \frac{\sqrt{2L}}{n\pi} \right) \cos \omega_0 t$$

$$\text{b) } -A \cos \omega_0 t \frac{\sin[\omega_0(x - L)/s]}{\sin(\omega_0 L/s)}$$

- 3.17 b) Use a Taylor series to expand  $\sin[Lk_n + L(k - k_n)]$ .

## Chapter 4

- 4.1  $E \approx 25$  V/m.  
4.3 b)  $C = (A\epsilon_0/d)$  for a parallel plate capacitor. Suppose that the surface around which we integrate  $\mathbf{B}$  in (4.4), rather than being a plane perpendicular to the wire and pierced by it, is deformed to pass between the plates of the capacitor, with now no physical current threading it anywhere (Fig. 4.10).  
4.4 b) See vector identities in text.

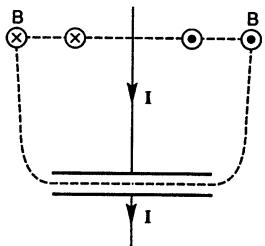
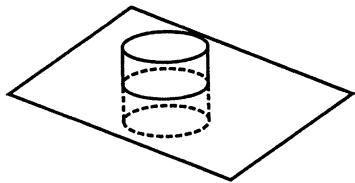


Fig. 4.10. Problem 4.3

- 4.6 a) Drop a vector  $\mathbf{V}(t)$  from the origin unto the plane.  
b) A vector field with zero curl will have as Fourier components harmonic waves pointing along the direction of propagation.  
4.7 a) See (4.13, 15).

4.9  $E_0 \sim 10^3$  V/m.

4.10 See Fig. 4.11.



**Fig. 4.11.** Problem 4.10

4.11 There are two directions of polarization for every  $\mathbf{k}$ .

4.12 b)  $\mathbf{J} = (ie^{-i\omega t})/(\mu_0\omega)(k_1 k_3/k)(\sin \pi x/a)\mathbf{j}$ . Check with Fig. 4.5.

4.13 a)  $\nabla \times (\nabla \times \mathbf{E}) = \nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E}$ .

4.14 c)  $u(r) = a/r$ .

f)  $z_0 = |n(b/a)/(2\pi\epsilon_0 c)$

4.15 a) The new “charge density” becomes

$$\varrho + \epsilon_0 \partial/\partial t \nabla^2 \chi ,$$

and the new scalar potential

$$\phi + \partial \chi / \partial t .$$

$$4.15 \quad b) \nabla^2 \chi = -\frac{1}{c^2} \frac{\partial \phi}{\partial t} .$$

4.18 a) Let  $\mathbf{j}(\mathbf{r}') = I\delta(\mathbf{r}')/2\pi r'$  so that

$$\int \mathbf{j}(\mathbf{r}') \cdot d\mathbf{S} = \int \int \mathbf{j}(\mathbf{r}') r' dr' d\phi = ?$$

b) Which of the variables  $r, \phi, z$ , does  $\mathbf{A}$  depend on?

$$c) \mathbf{B}(\mathbf{r}, 0, \phi) = (\mu_0 I/2\pi r)(L/\sqrt{L^2 + 4r^2})\mathbf{a}_\phi .$$

4.21 a) The driven wave equation for  $\mathbf{A}$  may be solved by the same procedure as the driven wave equation for  $y$  in Chap. 3. Note that

$$\int_0^d \int_0^b \int_0^a \mathbf{u}_{TE_{103}} \cdot \mathbf{u}_s dx dy dz = \frac{V k_1^2}{4k^2} \delta_{TE_{103},s}$$

where  $V = a bd$  is the volume of the cavity.

b) See Chap. 3 for resonance in distributed systems.

$$c) V = - \int_0^b E_y dy .$$

$$d) C = \epsilon_0 V / 2b^2 \equiv \epsilon_0 A / 2b \quad LC = 1/c^2 k_{103}^2 \equiv 1/\omega_{103}^2 \quad (c \text{ is the speed of light}).$$

$$4.22 \quad \mathbf{A}_1(\mathbf{r}, t) = \frac{-\mu_0 p_0}{4\pi r} \omega \sin(\omega t - kr) \mathbf{a}_z .$$

For  $kr \gg 1$ ,  $k = \frac{2\pi\omega}{c}$ ,

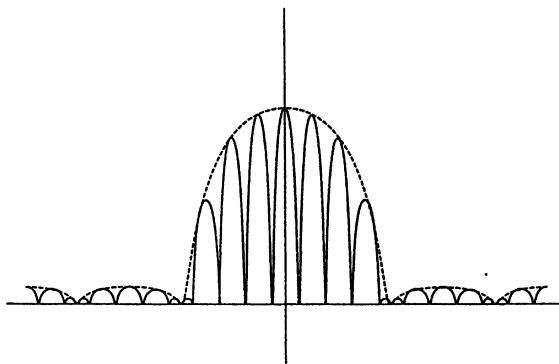
$$\mathbf{B}(\mathbf{r}, t) \approx \frac{\mu_0 p_0}{4\pi r} \omega k \cos(\omega t - kr) \sin \theta \mathbf{a}_\phi$$

$$\nabla \times \mathbf{B}(\mathbf{r}, t) \approx \frac{\mu_0 p_0}{4\pi r} \omega k^2 \sin(\omega t - kr) \sin \theta \mathbf{a}_\theta$$

4.A.1 Some lines end at the charges in the box.

## Chapter 5

- 5.2 b)  $\sim 6$
- 5.6  $\lambda \ll a$ ; two infinite half planes.
- 5.7 Can the integral of the phase factor be written as the product of two independent integrals? What happens to the pattern as we go to the long horizontal slit?
- 5.9 b)  $d \simeq 0.11$  mm. Total width  $\simeq 3.6$  cm.
- 5.10 b) Fig. 5.12.



**Fig. 5.12.** Problem 5.10b. Interference fringes modulated by the diffraction envelope;  $d \simeq (9/2)a$

- 5.17 c)  $\delta = 6.7 \times 10^{-4}$  mm

$$\text{d) } \nabla \times \mathbf{B} = \mu_0 \left[ \frac{\partial \mathbf{P}}{\partial t} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right],$$

if all current is viewed as polarization current. In case of a magnetization  $\mathbf{M}$ ,  $\mathbf{B}$  can be replaced by  $\mu_0 \mathbf{H} = \mathbf{B} - \mu_0 \mathbf{M}$  in the Maxwell equation.

## Chapter 6

- 6.7 c) If  $A$  and  $B$  do not commute a particle can be at most in the eigenstate of one of the operators, with a statistical spread for the observable

belonging to the other operator. If the operators do commute, there is the possibility of finding an exact value for both observables; i.e., it is possible within the wave mechanical framework to describe the situation where any spread for both variables simultaneously vanishes. It is not surprising that in this case the measurement of both quantities can in fact be simultaneously carried out without disturbing the state of the particle; the measurements are said to be *compatible*.

- 6.8 Earth-moon distance =  $3.82 \times 10^8$  m. Mass of earth =  $7.36 \times 10^{22}$  kg.
- 6.9 a) See Example 6.3 for  $x > a, x < -a$ . See Problem 6.11 for boundary conditions.  
 b)  $R^2 u_n(x) = u_n(x)$ . What does this say about the eigenvalues of the operator  $R^2$ ?
- 6.12 Consider  $\sqrt{P}$  in light of Schroedinger's time-independent equation.
- 6.14  $|A_0(t)|^2 = 1 - \sin^2(Wt/\hbar)$   
 $|A_1(t)|^2 = \sin^2(Wt/\hbar)$   
 See Problem 3.9 to complete your discussion.
- 6.19  $I_m \sigma_{21} = \frac{(\mu E_0/2\hbar)\tau(\Delta N_0/N)}{1 + (\omega - \omega_0)^2\tau^2 + (\mu E_0\tau/\hbar)^2}$   
 $\chi_2(\omega) = \frac{\mu^2\tau(\Delta N)_0}{\varepsilon_0\hbar} \frac{1}{1 + (\omega - \omega_0)^2\tau^2}$   
 $\chi_1(\omega) = \frac{\mu^2\tau(\Delta N)_0}{\varepsilon_0\hbar} \frac{(\omega - \omega_0)\tau}{1 + (\omega - \omega_0)^2\tau^2}$   
 when  $(\mu_0\varepsilon_0/\hbar)^2\tau^2 \ll 1$ .

$$(\Delta N)_0 = N(\varrho_{11} - \varrho_{22})_0$$

Plot  $\chi_1(\omega), \chi_2(\omega)$  as functions of  $\omega$  and notice the resonance at  $\omega = \omega_0$ . Temperature, for instance, will keep  $(\varrho_{22})_0 \neq 0$ . For more discussion refer to Chap. 8 of the text by Amnon Yariv, listed under Further Reading.

- 6.20 1 part in  $5 \times 10^9$ . Herzberg confirmed the effect by measuring the actual line to an accuracy of 1 part in  $10^6$ .

## Chapter 7

- 7.6 a)  $\Delta m = \varrho A \Delta l$ ; integrate over  $\Delta l$ .  
 b)  $p(l)/\varrho + gy + 1/2u^2(l) = c$
- 7.7 c) ii) If  $y_1$  and  $y_2$  are the same function of their respective compound arguments, what can you say about  $\partial[y_1(x, t) + y_2(x, t)]/\partial t|_{t=0}$ ? See also Problem 3.1.  
 c) iii) To find  $P(x, 0)$  we must know  $p(x, 0)$  in the interval. Subsequently, we can find  $p(x, t)$  in terms of  $u(x, t), u(x, 0), P(x, 0)$  along the characteristic. Two intersecting characteristics will always yield  $p(x, t), u(x, t)$ .

- 7.11 b) If  $f_2$  is large  $v(x, t)$  becomes a function of  $\Lambda f_1$ . Note the symmetric form of  $v(x, t)$  in (7.43).  
 c,d) We can take  $|t|$  as large as we please.  
 e) for  $t \rightarrow -\infty$  Soliton 1 leads; for  $t \rightarrow +\infty$  Soliton 2 leads. Soliton 2, which has the larger velocity, given by  $\alpha_2^2$ , has passed Soliton 1. However, the shapes of the two solitons have not changed in view of the constant functional dependence of  $v(x, t)$  on  $x$  and  $t$  for any given soliton  $j$ ,  $j = 1, 2 : f_j/(1 + f_j)^2$ . A change of phase for each soliton can be identified.

- 7.12 See Problem 1.6a.

$$\begin{aligned} A_n &= 0 & B_n &= \gamma_n(t) & n &= 0, 1, 2 \dots \\ A(k) &= \frac{1}{\sqrt{2\pi}} & B(k) &= \frac{\beta^*(k)}{\sqrt{2\pi}} & k > 0 \\ A(k) &= 0 & B(k) &= \frac{\alpha^*(k)}{\sqrt{2\pi}} & k < 0 \end{aligned}$$

with  $\omega_k > 0$  for all  $k$ .

$$\begin{aligned} 7.14 \quad \text{Im}\{\beta\} \int_0^\infty e^{-x^2/4n} \sin(2kx) dx &= -\text{Im}\{\beta\} \int_{-\infty}^0 e^{-x^2/4n} \sin(-2kx) dx \\ &\quad \int_0^\infty e^{-x^2/4n} \beta^*(k) e^{-2ikx} dx = \int_{-\infty}^0 e^{-x^2/4n} \beta(k) e^{2ikx} dx \end{aligned}$$

the normalization of a scattering state.

$$\begin{aligned} 7.15 \quad \text{a) } \frac{\partial}{\partial x} (u_x R - u R_x) &= u_{xx}(u_t - 2vu_x - 4\omega^2 u_x) - u(v_{xxx}u - 4v_x u_{xx}) \\ &\quad - u(vu_t - \omega^2 u_t - \omega_t^2 u + v_t u) \\ &\quad + u(2v + 4\omega^2)(v_x u - \omega^2 u_x + vu_x), \end{aligned}$$

which simplifies if one uses (7.46), or see Drazin and Johnson, loc. cit. p. 68.

$$7.18 \quad \text{c) } h(x, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty e^{(ik^3 t + ikx)} dk.$$

- 7.19 b) Yes; however, direct substitution into the KdV equation to check is not recommended.

$$7.20 \quad \text{b) } B(x, t) = \frac{-1}{2\pi} \int_{-\infty}^\infty \frac{\exp [i(kx + 8k^3 t)]}{1 + ik/K} dk.$$

If we let  $k = y(x/t)^{1/2}$ ,

$$B = \frac{-2K}{2\pi} \int_{-\infty}^\infty \frac{\exp [i(x^3/t)^{1/2}(y + 8y^3)] dy}{2K(t/x)^{1/2} + 2iy}$$

leading, as it turns out, to an oscillatory dispersive wave to the left, decaying as  $t^{-1/3}$ , as  $t \rightarrow \infty$ .

# Index

- admittance 83
- Ampere's law 88
- amplifier 46
- anode 127
  
- basis functions 2, 167
  - complete orthonormal set 63
  - complete set 2, 68, 92, 97
  - complete, periodic set 80
  - continuous basis 167
  - discrete basis 167
  - function space 2
  - normalization 167
  - normalized 3
  - orthogonal 3
  - orthonormality 3, 14
- beam 144
  - of particles 171
- Bernoulli's equation 241
- Bohr, N. 130, 159
  - Bohr radius 206
- Born, M. 161
- boundary conditions 61, 75, 111, 145, 163
  - for the electric and magnetic fields 99
  - homogeneous 71
  - on string 59
  - periodic ~ 171, 206
- de Broglie, L. 158
  - de Broglie relation 158
- Byron, F.W. Jr. 17, 31
  
- capacitance 98
- capacitor
  - parallel plate ~ 117
- cathode 127
- Cauchy's Theorem 9
- Cellular Processor 262
- complete set 29
- completeness 81
- complex exponential 2
  
- complex numbers 7, 38
  - de Moivre's Theorem 25
- Compton, A.H. 158
- condensed matter physics 191
- correlation function 28
- correspondence principle 130
- Coulomb's law 123
- coupled oscillator 80
- curl 88, 89, 103, 125
  - angular velocity 118
- circulation 90
  
- damped oscillations
  - damped traveling waves 84
- Davisson, C. 158
- Debye-Scherrer 141
- degrees of freedom 2, 55
- dielectric constant
  - complex ~ 156
- dielectrics 142
  - dielectric constant 142
- differential equation
  - linear 35
  - separation of variables 206
- differential equations 17
  - Korteweg-de Vries equation 229
  - Korteweg-de Vries pde 233
  - separation of variables 241
- diffraction 129, 133–139, 147
  - *m*-th order spectrum 153
  - circular aperture 139
  - Debye-Scherrer ~ pattern 154
  - of electron beams 159
  - Fraunhofer ~ 136
  - Fresnel ~ 136
  - Fresnel zone 152, 153
  - grating 138, 139, 153
  - Laue ~ pattern 154
  - pattern 134, 137
  - phasors 154
  - resolving power 153
  - single slit ~ 153

- sodium doublet 153
- X-ray  $\sim$  139
- diffusion equation 28
- dipole radiation 107–110
- Dirac delta function 3, 9, 22, 23
  - potential 199
- Dirac, P.A.M. 3
- Dirac equation 191
- dispersion
- group velocity 221–223
- displacement field 142
- distributed systems
- electromagnetic  $\sim$  98
- divergence 88, 89, 103, 124
- driven motion
- driven behavior 98
  
- eigenfunction 162
- eigenfunctions
  - completeness 167
  - expansion in  $\sim$  of  $H$  168
  - normalization condition 170
  - of a quantum mechanical operator 167
- eigenvalue
  - continuous  $\sim$ s 169
  - energy  $\sim$ s 164
- Einstein, A. 130, 265
- photoelectric effect 151
- special relativity 191
- electric inductance 34
- electric susceptibility 141
- electric transmission line 35
- electrical capacitance 35
- electrical resistance 35
- electrical susceptibility 156
- electromagnetic cavities 98–102
- electromagnetic cavity 97, 158
- electromagnetism 87, 110
  - Biot-Savart law 122
  - coaxial cable 120
  - energy densities 96
  - fields 106
- electron microscopes 159
- electrostatic potential 110, 114, 121
- electrostatic potential  $\phi$  105
- energy bands 130
- energy levels 130
- escapement 47
- even symmetry
  - even function 26
- expectation value 169
- expected mean 166
  
- Faraday's law 88
- Feshbach, H. 73, 85
- Feynman, R.P. 98, 127, 196, 215
- fine structure 191
- flux 134
- forced motion 33, 41–46, 67
  - driving terms 87
  - of a string 71
- Fourier analysis 3, 11–17, 178, 272
  - cosine Fourier transform 132
  - Faltungstheorem 27, 52
  - Fourier cosine representation 20
  - Fourier representation 16
  - Fourier transform 3, 12
  - Fourier transforms 4
  - FT of  $\delta(x)$  13
  - Gibbs' phenomenon 21
  - harmonic components 3
  - Parseval's theorem 27
- Fourier, J. 3
- fractals 256
  - fractal fingers 258
- Fredholm integral equation 237
- Fuller, R.W. 17, 31
  
- $\gamma$  particle 196
- $\gamma$ -rays 129
- Gaponov-Grekhov, A.V. 250
- Gauss's law 88, 104
- Gauss's theorem 90, 111, 148, 154
- Gauss, F. 5
- gaussians 5
- Geiger counter 179
- Gelfand-Levitant-Marchenko equation 238
- generalized functions 4, 10, 22, 23
- geometric optics 144
- Germer, L. 158
- good functions 4, 5, 23, 82
  - fairly 24
- gradient 88, 89, 124
- Green's function method 33, 44, 71, 72, 110, 189
  - expansion of the Green's function in eigenfunctions 76
  - free particle Green's functions, 197
  - Green's function for a stretched string 77
  - harmonic oscillator 181
  - magic rule 73, 83, 147
  - principle of reciprocity 150
  - quantum field theory 193
  - quantum Green's function 195

- retarded Green's function 193
- Green's theorem
- one dimensional equivalent 72
- Green, G. 44
- Green's theorem 148
- Gutzwiller, M.C. 265, 275
- Hamilton, W.R. 162
- harmonic oscillator 182
- Heisenberg uncertainty principle 133, 171–179, 194
- Heisenberg, W. 161, 180
- Helmholtz equation 148
- Helmholtz theorem 104
- Hermetian operator 187, 211
- Hilbert space 17–18
- Hilbert, D. 2
- Hooke's law 35
- Huygens' principle 135, 152
- hyperfine structure 192
- idealization 57, 99
- impedance
  - characteristic  $\sim$  of cable 120
  - complex  $\sim Z$  for a mode 123
  - complex  $\sim Z$  for an idealized circuit 53
  - complex  $\sim$  of the string 83
  - complex mechanical 42
  - mechanical 82
- inductance 98
  - of a long solenoid 50
- initial conditions 62
- initial value problem 64
- interference 129, 147
  - patterns 138
- intrinsic magnetic moment 191
- inverse scattering 233–241
- Jordan, P. 161
- Kay, I. 233
- Kronecker delta 162
- Lamb shift 214
- Lamb, W.E. 198
- lasers 133
- von Laue, M. 139, 154
  - spot 141
- Lenz's law 34
- Lighthill, M.J. 5, 31
- linear systems
  - linear differential equations 36
- Liouville's theorem 84
- longitudinal field 104
- Lorentz gauge 121
- magnetic imaging 205
- matrix mechanics 161
- Maxwell equations 155, 192
- Maxwell's equations 88
  - differential form 88, 91
  - integral form 88
- Maxwell, J.C. 95
- Michelson, A.A. 95
- Morse, P.M. 73, 85
- Moses, H. 233
- natural motion 33, 205
  - natural behavior 98
  - natural frequency 39
  - of systems 36
- normal modes 60, 65, 80, 83, 158, 163
  - classification 100
  - eigenstates 191
  - electric field in the mode 116
  - in classical physics 174
  - longitudinal 81
  - modes of polarization 96
  - TE<sub>103</sub> mode 119
  - transverse 81
- nuclear physics 192
- obliquity factor 135, 151
- observable 162, 166
- odd symmetry
  - odd function 26
- operator 162, 166
  - Hermetian 164
- optical disturbance 135
- oscillator
  - harmonic  $\sim$  192
  - oscillators 33, 44, 46
  - damped oscillations 37
  - damping 33, 190
  - free oscillations 39
  - half-power point 49
  - response 43
- overlap integral 4, 7, 29
- Panofsky, W.K.H. 103, 127
- particles 130
- Pauli, W. 192
  - Pauli exclusion principle 192
- periodic table 191
- phase 42, 51
  - unknown 83
- phasors 153

- Phillips, M. 103, 127  
 photoelectric effect 130  
 photon 130  
 physical optics 129  
 – crisis in optics 157  
 Planck, M. 130, 158  
 – Planck's constant 130, 158  
 Poisson equation 110  
 Poisson, S.D. 110, 111  
 polarization 97  
 – plane-polarized 96  
 polarization charge 154  
 polarization of materials 141  
 positron 192  
 Poynting vector 96  
 Poynting, J.H. 96  
 prisms 174  
 projection formula 18  
 quality factor  $Q$  33, 40, 47  
 – for a cavity 102  
 Quantum Chaos 265
- Rabinovich, M.I. 250  
 radiation  
 – black body  $\sim$  158  
 – coherence length 135  
 – electromagnetic  $\sim$  129  
 – incoherent  $\sim$  133  
 – intensity 137  
 radiation zone 123  
 Randomness 251–253  
 Rayleigh, Baron 138  
 – Rayleigh criterion 138  
 reactance  
 – mechanical 42  
 refraction 141, 143  
 – complex index of  $\sim$  155  
 – index of  $\sim$  144  
 – refractive index 144  
 – Snell's law 155  
 resonance phenomenon 33, 44, 81, 123  
 – condition of resonance 70  
 – resonance frequency 98  
 – resonance response 102, 205  
 – vicinity of resonance 51  
 retarded potential 107  
 retarded solution 106  
 root square deviation 166  
 Russell, J.S. 230  
 Rydberg atom 268
- Schroedinger equation 160  
 Schroedinger, E. 158, 191
- Schwartz, L. 5  
 second quantization 192  
 skin depth 155  
 solitons 229–241, 245  
 sound wave 81  
 source 111  
 – antenna 106  
 – dipole antenna 106, 107, 123  
 – source fields 103  
 sources 87  
 square wave function 59  
 square wave function  $F(x)$  18  
 stationary state 169  
 statistical mechanics 158  
 Stokes's theorem 91  
 structures  
 – disorder 250  
 – ordered structures 250  
 superconductivity 191  
 symmetry  
 – even 7  
 – odd 7  
 systems  
 – distributed 35, 87, 98  
 – linear mechanical 35  
 – lumped 35, 55
- Temple, G. 5  
 Thomson, G. 159  
 time constant 40  
 topographical vortices 260  
 transverse field 104
- unit vectors  $i, j, k$  3
- vacuum diode 114, 126  
 – Child–Langmuir law 115  
 vector potential 103, 104  
 viscous drag 35
- wave function 159  
 wave mechanics  
 – commuting operators 207  
 – ground state 183  
 – Heisenberg operators 193  
 – operators 193  
 – pair creation 196  
 – postulates of  $\sim$  161–171  
 – quantum electrodynamics 196  
 – raising and lowering operators 184  
 – reflection operator 208  
 – Schroedinger's cat 180  
 – stationary states 270

- time dependent wave equation 185
- time-dependent problems 211
- tunnel 200
- two-level system 212
- wave-particle duality 179
- wave mechanics operators
  - annihilation and creation operators 192
- waves
  - amplitudes 3
  - angular frequency 2
  - dispersion 174
  - fundamental frequency 66
  - group velocities 241
  - group velocity 144
  - nonlinear  $\sim$  223–229
  - packet 144, 178
  - period 2
  - phase shift 3, 39
  - phase speed 174
  - phase velocity 144
  - plane solutions 92
  - polarization 135
  - scattering states 235
  - spherical 97
  - standing 99
  - surface  $\sim$  218
  - transverse 92
  - transverse displacement 56
  - transverse electric 101
  - transverse field 92
  - transverse magnetic 101
  - traveling 66, 82, 99, 115
  - on water 217
  - wave motion 58
  - wave number 2
  - wave-particle duality 157
  - wavelength 2
- X-rays 158